DOCUMENT RESUME

ED 150 158 95 TM 006 781

AUTHOR Crawford, W. John; And Others
TITLE Texas Teacher Effectiveness Project: Stability Correlations Between First and Second Year Data.
INSTITUTION Texas Univ., Austin. Research and Development Center for Teacher Education.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO UTR&D-R-75-14
PUB DATE Jul 75
CONTRACT NIE-C-74-0089; OEC-6-10-108
NOTE 149p.; For related documents, see ED 091 394, 099 345, and TM 006 784 ; Not available in hard copy due to print quality

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Academic Achievement; Achievement Gains; Behavior Rating Scales; Class Management; *Classroom Observation Techniques; *Classroom Research; *Correlation; Effective Teaching; Elementary School Students; Elementary School Teachers; Grade 2; Grade 3; Interaction Process Analysis; Primary Education; Questioning Techniques; Reinforcement; Research Problems; Student Behavior; *Student Teacher Relationship; Tables (Data); *Teacher Behavior; Teaching Techniques; *Test Reliability
IDENTIFIERS Classroom Environment Scale; Dyadic Interaction Analysis; Teacher Effectiveness Study; Texas

ABSTRACT

Pearson correlation coefficients between first and second year classroom process data from the Texas Teacher Effectiveness Project are presented. These data are from 19 second and third grade teachers whose classrooms were observed during two consecutive school years. The stability of classroom behaviors varied greatly across the two years. Factors which affected stability included: frequency of the behaviors, reliability of the instruments, teacher vs. student control of the behaviors, and the amount of structure afforded by the contexts within which behaviors occurred. Possible explanations for the observed differences in stability of these process variables (both high and low inference) are offered, and implications of these data for classroom research are discussed. (Author/MV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

4025

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
CERIC
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED150158

TM006 781

The Research and Development Center for Teacher Education

UTR&D

The University of Texas at Austin

Texas Teacher Effectiveness Project:
Stability Correlations Between
First and Second Year Data

W. John Crawford,

Jere E. Brophy,

and

Carolyn M. Evertson

Report 75-14
The University of Texas at Austin
July, 1975

This project was supported by the National Institute of Education Contract OEC 6-10-108, Research and Development Center for Teacher Education, and by Contract NIE-C-74-0089, Correlates of Effective Teaching. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that office should be inferred.

Abstract

Pearson correlation coefficients between first and second year classroom process data from a two-year replicated study are presented. These data are from 19 second and third grade teachers whose classrooms were coded during two consecutive school years. There was great variation in the degree to which classroom behaviors were stable across years. Factors which affected stability included: frequency of occurrence of the behaviors; reliability of the instruments; teacher (vs. student) "control" of the behaviors; and the amount of structure afforded by the contexts within which behaviors occurred. Possible explanations for the observed differences in stability of these process variables (both high and low inference) are offered, along with discussion of the implications of these data for classroom research.

The Texas Teacher Effectiveness Project:

Stability Correlations between First and Second Year Data

This report presents results from the Texas Teacher Effectiveness Project (Brophy and Evertson, Note 1) concerning stability across years, in correlations of classroom process variables. The process data were collected by trained observers in second-and third-grade classrooms. The study was replicated during two school years: 1971-1972 and 1972-1973. Data collection on the same variables during two separate years permitted correlation of the data sets to assess stability of process behaviors. By examining these Pearson product-moment correlations, we were able to see which variables were significantly stable across both years of the study and which ones were not.

Though the data reported here consist of correlations of process variables between "Time 1" (the first year) and "Time 2" (the second year), it should be noted that they are <u>not</u> equivalent to test-retest reliability information on the instruments employed in the study. More accurately, the data reflect the degrees of consistency of behaviors coded in the various categories of teacher-student interaction. To obtain data parallel to test-retest reliabilities, we would have had to collect and correlate data on the same teachers and students in the same contexts at two different times. Even this procedure would not have yielded totally clean measures of instrument reliability, since the same teachers and students would not be likely to interact in the same ways at different times in the school year. Also, in our replicated study, the student populations changed from

the first year to the next, so that the same groupings of teachers and students would not be observed in both years; this factor will be discussed more thoroughly later on in the report. Inter-coder reliabilities were computed for the instruments; these are presented in the tables either as intra-class correlations, percentage agreement figures, or "category-usage" information. These data show the degree to which the instruments were used consistently by different coders.

Thus, the findings discussed here primarily are indications of which types of teacher-student interactions and teacher behaviors were stable across two years for this sample of teachers. Discussion of factors which affect the degree of stability follows later. Since the stability correlations are part of a larger process-product correlational study, we first will give some background information and references, so that readers will understand the whole study and the context in which the present data must be considered.

Sample Selection

The Texas Teacher Effectiveness Project investigated process-product relationships between high and low inference process measures from classroom observations and mean residual achievement gain scores on the Metropolitan Achievement Test. Instead of selecting teachers randomly, the sample of teachers was chosen on the basis of their previously established records of consistency in producing student learning gains. MAT scores for three consecutive years were obtained for 165 second and third grade teachers, for each of the five subtests for which data were available.

Each student's raw gain score (grade level equivalent) was converted to a residual gain score for each subtest. These residual gain scores then were grouped by class, and mean subtest residual gain scores were computed for each of the 165 teachers for each of the three years of data. It then was possible to select from the total of 165 a smaller sample of teachers who were notably consistent across subtests, sexes, and time (the previous three-year period) in the relative achievement gains of their respective students (Brophy, 1973; Veldman and Brophy, 1974).

It was felt that this sample was particularly appropriate for a process-product study investigating correlates of effective teaching, since teachers with records of consistency in producing student learning gains would be more likely to show consistent classroom behavior. Following the suggestions of Soar (1972), teachers were selected from the full range of effectiveness scores, rather than only from the extremes. This yielded approximately normal distributions of effectiveness scores for the observed teachers.

Method

Thirty-one teachers participated in the first year of the study. Of these, 12 were not available for participation in the second year due to leaves, transfers to new grades, and a few refusals. However, nine other consistent teachers were added the second year, for an N of 28. In any case, the "common sample" included 19 teachers; i.e., there were 19 teachers who participated in both years of the study. The stability coefficients reported here, then, are measures of the consistency of the process variables

measured in 19 classrooms across two consecutive years, and the maximum possible N for these data is 19.

In the first year, four observations (about 10 hours) were made in each classroom, split evenly between mornings and afternoons when possible. In the second year, 14 observations (about 30 hours) were taken. Coders worked in pairs in Year 1, and their codes were averaged to produce final scores. In Year 2, coders worked in pairs until they reached an 80% agreement criterion, then coded individually (to double the number of observations). To counteract systematic coder differences, each teacher was coded (roughly alternately) by two different coders in Year 2. The high inference measures (Evertson and Brophy, 1973) included ratings and checklists. Some were used during each classroom visit, but most were filled out only twice; once at mid-year, and once at the end of the year. The low inference behavioral data were collected with an expanded version of the Dyadic Interaction Observation Coding System (Brophy and Good, 1970). With this system, teacher-student interactions are coded in many categories, some of high and some of low frequency (see Brophy and Evertson, Note 2).

Some low stability in the low frequency categories is to be expected, due to less reliable behavioral samples. Also, the probabilities for significant correlations were limited by the low N (maximum N = 19). Another factor lowering the probabilities of significant stability coefficients was the fact that the first year data represent only 10 hours of classroom observation time. Since this is a relatively small amount, there is a greater likelihood that the resultant data are not representative of the teachers' and students' typical classroom behavior. These remarks preface

the presentation of data by indicating that variables which did correlate significantly positively across years are likely to reflect highly stable classroom behaviors, given the difficulty of obtaining significant r's.

One might expect the teachers in this sample to be more stable in their teaching behaviors than a random sample of teachers at the same grade levels, however, since they were selected for their consistency in producing student learning gains. Thus, the set of significantly correlating variables represents the most stable process behaviors of teachers who are highly consistent in producing product outcomes. This is important for considerations of external validity, since a random sample of teachers probably would show fewer (and perhaps different) stable behaviors.

In addition to being more consistent in producing student learning gains, our sample of teachers differed from a random sample in having more years of teaching experience and in being more "traditional" in their responses to a questionnaire (Sherman, Brophy, Crawford, and Evertson, in press). These factors could have influenced the number and types of variables which were significantly stable across time.

The present report, dealing with stability within contexts but across years, is one of two major investigations of stability in the process behaviors measured in this study. For a report of stability across different contexts (morning, whole group; afternoon, whole group; and reading group), but within years, see Brophy, Evertson, Crawford, King, and Senior (in press).

Data Analyses

There are eight subsets of data for which stability coefficients will be reported: classroom observation scales, behavior checklists, coder ratings, time utilization variables, percentage estimates, morning behavioral data (whole class), afternoon behavioral data (whole class), and small group (reading group) data. Readers interested in coding procedures, resultant variables, or any other details concerning the collection of data with these instruments should consult Brophy and Evertson, Note 1.

Also, it should be noted that, in addition to the previously mentioned problems which would tend to minimize correlations (low N, low frequency of occurrence, low variance, etc.), there is the problem of instrument reliability. It was almost always less than perfect, and in some cases was poor. No corrections for this attenuation were made on the correlations. Non-normal distributions (designated with an asterisk and/or a "NONE" in the data tables) were not normalized before correlations were computed. It was felt that avoiding these methods for increasing the likelihood of obtaining a significant r would make our results the most conservative estimates of stability.

While this procedure undoubtedly decreased the number of significant r's, it also gives us greater confidence in the variables which were significantly stable. The "valid data" N's range from 19 down to zero. The r's, N's, and p's are reported in the tables whenever available, but only variables with N's greater than or equal to five are included in the discussion of the results. The most common reason for "missing data" -- where N was less than 19 -- is that the variable never occurred in one or more of the classrooms.

The probability values for the correlations, given the valid N for a particular computation, are presented in the data tables. These probability values convey more information than a "significant-not significant" distinction by also showing non-significant trends. In this report, "significant" means a p value at or below .10. Sometimes "highly significant" or "highly stable" is used in referring to p values below .05 or .01.

The tables also indicate the modal score and the number of teachers who had scores different from these modal scores (N-MODE N). These data provide information about the location and degree of variance in the measures. Readers also will notice that some variables occur in more than one data subset, or even in different contexts within the same subset. This is because we collected "context-specific" data on many variables of import. For example, "teacher praise" was coded both for public response opportunity interactions and for private child created contacts. All of the low inference variables are repeated in three subsets (see tables 6, 7, and 8): morning whole class interactions; afternoon whole class interactions; and reading groups. This procedure was used to determine whether differential relationships existed among the same behaviors measured in different classroom situations (contexts).

The next sections of the report present the data. High inference ratings, checklists, and percentage estimates are discussed first, followed by the low inference data. The tables contain the stability correlations, valid N's, probability values, and distribution information on the first and second year data.

High Inference Measures

Emmer and Peck (1973) developed a set of high inference measures for coding various classroom behaviors. The twelve 5-point scales are based on factor analyses of five widely used low inference coding systems. Conversion to a high inference system allowed measures of these classroom variables to be obtained more quickly and easily. The measures include: student attention, mode of addressing questions, task orientation, pupil-to-pupil interaction, percent of time lecturing, positive and negative affect, lectures and explanations, clarity, teacher enthusiasm, convergent questioning, and student passivity.

For both years, the classroom observation scales showed mostly high inter-rater agreement. For the first year, the measure used was percentage agreement (within one point) between coders rating in the same classroom. For the second year, the inter-rater agreement figures were average intra-class correlations. The Year 1 percent agreement figures ranged from 72% to 95% on the 12 scales, averaging 83%. The Year 2 intra-class correlations between coders ranged from .12 to .89, averaging .58. In Year 2, one coefficient was .12 and another was .48; all others were above .60. The two scales with low reliability both dealt with level of classroom discourse. Since this was almost uniformly low in these early elementary classrooms (Brophy and Evertson, Note 1), the low agreement coefficients probably reflect low variability in the behaviors of interest.

Nine of the twelve classroom observation scales were significantly stable across years, with stability coefficients ranging from .55 on the convergent evaluative interaction scale to .88 on the task orientation scale (see Table

1). The three scales which were not significantly stable across years were pupil-to-pupil interaction, percent teacher presentation, and pupil passivity. One possible explanation for the lack of stability in these scales is that, in the second year of the study, the school district emphasized individualized teaching through the use of team teaching, learning centers, and other methods. This probably increased pupil-to-pupil interaction, decreased percent teacher presentation, and perhaps decreased pupil passivity in some schools (though not all), thereby producing a lack of stability for these three scales. Also, two of the three unstable variables are measures of student, not teacher, behaviors; the student populations in the study changed from one year to the next. Higher stability was expected for measures of teacher behavior than for measures of student behavior.

Insert Table 1 About Here

In general, we see that most of the Classroom Observation Scales are highly stable across time and reliably coded by observers. Most of these stable process variables which are controlled by the teacher and which are easily coded, have good data distributions, etc. and measure general classroom "mood" and "atmosphere." Also, they were taken several times during each observation (as opposed to the other high inference measures, many of which were designed to pick up specific instances of noteworthy classroom interactions or behaviors). Thus, the Classroom Observation Scales were averaged across many more samplings than the other high inference ratings and checklists.

Another set of high inference measures taken by observers in the classroom were checklist variables concerning whether or not the teacher used particular methods of handling problems. A low percentage of the stability coefficients for these variables reached the .10 or .05 significance levels, primarily because of missing data and generally low frequencies in most of the categories. These checklists were measures of specifically observed teacher behaviors. The coder had to actually see the teacher performing one of these activities before it could be coded. This necessarily lowers the probability of occurrence of significant correlations, since what is being recorded is not inference based on continuing general class activity, but rather, certain specific behaviors which typically occur infrequently.

Twelve dimensions were coded in this checklist subset, each of which had various sub-variables within it which the coder noted as being either present or absent in the teachers' repertoire of classroom behaviors. The inter-rater agreement figures were generally low; here, again, due to the fact if a specific behavior occurred, the chances were that both raters, if they viewed the teacher at different times, would not see the same behavior, thus resulting in low inter-rater agreement. The primary consideration in this subset though, is the low frequency of occurrence of most of the behaviors. Variables were not coded on a continuum, but rather the observers were trained to make a check if a behavior was present and not make a check if it was absent. This contributed to low variance and low frequency. The twelve categories of variables in this subset are: catch-up work, restriction of physical movement, punishment, rewards, appropriate assignments, distractions, student attitude, free-time available material, free-time

material used, use of peer tutoring, use of homework, and teacher underreactions.

Insert Table 2 About Here

As Table 2 indicates, the variables which were shown to be stable across years were variables which a coder would be likely to observe in a classroom in any one or two observations. They included such variables as: children having to get permission to leave their seats, scolding, classmates clapping or cheering, students seem amused by teacher, students seem to respect the teacher, available free-time materials (whether they be exhibits, aquariums, or the like), use of peer tutoring, assignment of homework, and teacher underreactions to control problems. These kinds of variables could be expected to occur every day, and therefore the coders would be much more likely to come in contact with these behaviors and be able to code them. In summary, then, the data from this subset do continue to conform to the pattern of maintaining high stability coefficients in the cases where expected, i.e. high frequency, commonly observed categories with high inter-rater agreement figures.

Another set of high inference data was obtained by having coders rate with five- and thirteen-point scales the degree to which certain teacher behaviors occurred. The variables that were coded on thirteen-point scales were affectionateness level, most intense positive affect observed, and most intense negative affect observed. All other variables were coded on five-point scales.

For these data, percentage agreement figures between coders were computed on the basis of raters agreeing within two points on the thirteen-point

variables and within one point on all other variables. The inter-rater agreement figures are generally high, ranging from 50% to 98% in Year 1, and from 56% to 94% in Year 2. These inter-rater agreement figures averaged 82% for the 42 variables for the Year 1 data, and 80% for the 42 variables for the Year 2 data. The average inter-coder reliabilities for the Year 1 data were about 82% for both the significantly stable correlations and the ones which were not significantly stable. The variables which yielded significant correlations across time did have higher coder reliability figures than the nonsignificant variables for Year 2: 82% versus 75%. While this difference is not statistically significant, it does indicate, as would be expected, that the stable variables were also coded more reliably by coders in the second year of this study.

Insert Table 3 About Here

With the exception of certain variables expected to be affected by the school district's second year emphasis on individualization, the variables in this subset do present a picture of being relatively stable measures across time. This fact combined with high inter-coder reliability figures should make these ratings good measures of classroom behavior. Lack of stability in some of these variables could possibly be explained by the fact that the school district's mandated curriculum changes which went into effect in the second year of the study did include much more stress on individualization and less "classroom discussion-response opportunity' type of instruction. Especially the variables "speaks with class versus individuals," "trial and correction versus preparatory instruction," and "quantity of directions" could be expected to be affected by this factor.

A fourth set of high inference data was collected by having the observers note percentages of time that teachers devoted to dealing with several different classroom situations. The stability correlations for these variables are presented in Table 4. Three time utilization variables were significantly stable across years: "proportion of time spent in art,"

Insert Table 4 About Here

"proportion of time spent in spelling," and "proportion of time spent in transitions." Two other variables in this subset, "proportion of time spent in language arts" and "proportion of time spent in morning routine," had significantly negative stability correlations across years. The most likely explanation for the negative correlation for the "proportion of time spent in language arts" variable is that three of the teachers in the common sample taught no language arts the second year, though they did instruct in language arts in the first year of the study. Also, another teacher was coded for no language arts teaching in Year 2 because she taught language arts in the afternoon, but never was observed in the afternoon. Thus, we see that four teachers (over 20% of our common sample) showed considerable percentages of time spent in langauge arts teaching for Year 1, while being coded for zero percentages of their time spent in this activity for Year 2 of the study. This fact, combined with fluctuations in other teachers' time spent in this activity (due to team teaching, primarily), produced the negative correlation.

The other negative stability coefficient, "time spent in morning routine," was probably caused by the restricted range of scores for all the subjects: from zero percent of total time to only about 4%. This lack of meaningful

variance produced genuine instability in this variable, and any sizeable change at all in the teacher's habits of amount of time spent in morning routine could have produced the negative correlation.

Other variables in this subset which showed a lack of stability were total available time, total teacher controlled time, proportion of teacher controlled time over available time, proportion of time spent in math, in reading group, in social studies, and proportion of time spent in special activities. In general, this subset seems to lack stability, indicating that the teachers who were in our common sample used proportionately different amounts of their time for different activities in the first year of the study as compared with what they did in the second year. It seems most likely that these changes across time are the result of the individualized instruction curriculum changes introduced in the second year of the study.

Another high inference data set collected was estimates of percentages of different methods teachers used in coping with various situations in the classroom. In the first year of the study, the percentage estimate ratings were made twice by the coders; after the third and after the fourth observations (there were only four observations taken in the first year of the study). In the second year of the study, two percentage estimate ratings were also made by coders; one after the sixth or seventh observation and one after the last observation (number of observation of teachers was usually 14 in Year 2).

There were five major variables within this subset, each of which has several sub-variables within it. The major variables are: methods used to call for attention, percent attending, what the teacher does when a child doesn't understand, locations of interactions, and what the teacher does when the child is "stuck" on a word in a reading group. Of the seven sub-variables

which had significant stability correlations across years, four were "calling attention" variables, one was "percent of class attending," and two were "child doesn't understand" variables. The stability of these variables would be expected because of their high frequency of occurrence and the high rater

Insert Table 5 About Here

agreement figures. Also, student attention was coded during each observation with the Classroom Observation Scales, which were discussed earlier. The high percentage of stable variables in this category is consonant with the fact that the coders are more aware of this behavior and probably can rate more accurately because of practice during each observation with the Classroom Observation Scales. Stability for any of the "calling attention" and "child doesn't understand" variables probably reflects stability in teacher classroom management and clarity.

The reading group variables in this subset generally have low stability. This could be because coders usually had a difficult time hearing the particular actions of the teachers in dealing with children in these groups. Usually coders were seated across the room from the group, and the teacher and children in the group attempted to keep their voices low so as not to disturb the rest of the class. While they were successful in this attempt, they also made it difficult for the coders to hear what was going on in the group. Another reason is the small number of observations for the first year data -- only four observations totally, and typically only two observations of reading instruction. Also, some new readers were introduced in the second year. Teachers were given the choice of three readers instead of only

being able to teach from the one set used the previous year. The "percent attending" variable is stable as might be expected; this is a behavior easily observed and coded (also, teachers are likely to be consistent in their demands for student attention).

Most of the "child doesn't understand" variables are not significantly stable across time. The most likely reason for this is the low frequency of most of the sub-variables in this category. This is especially true for the first year data, since each teacher was seen for a maximum of four observations only.

In general, the pattern previously established for the other high inference ratings continues to hold true for this subset also. The high frequency, easily observed variables tend to be stable across years, while the measures of behaviors which occur irregularly, have low inter-coder reliability, or are hard to code because of difficulty in hearing and/or seeing tend to lack stability as measured by correlation coefficients between the Year 1 and Year 2 data.

The other high inference subset with data for both Year 1 and Year 2 of the study is the "lesson presentation" variable subset. Here, data were collected on types of curricula used, methods employed, and amount of individualization of the teacher's instruction. Although the same classroom dimensions were measured in both years of the study, the data collection and sampling differ considerably between the first and second year, thus precluding any meaningful correlational analyses for these variables.

In the first year, the variables were measured with a low inference system used on a subsample of ten teachers, five of whom were high effective (i.e., produced high achievement gain scores in their students) and five of

whom were low effective (i.e., produced low achievement gain scores in their students). The data were taken continuously in low inference fashion in these ten teachers' classrooms.

In the second year of the study, data on these variables were collected in all 28 of the participating teachers' classrooms. Data collection methods in Year 2 were high inference five-point scales used for each of the dependent variables. Of the 28 classrooms in Year 2 of the study, only seven were coded for these data in Year 1. This maximum possible "valid data" N of 7 is too low for meaningful computation of stability correlation coefficients across years (no stability correlations will be presented in this report for these variables). The Year 2 data did yield meaningful relationships when correlated with the achievement test gain scores for the low SES and high SES groups (see Brophy and Evertson, Note 3), however.

Low Inference Measures

An expanded version of the Dyadic Interaction Observation System (Brophy and Good, 1970) was used to collect low inference data in both years of the study. Coders were extensively trained in the use of the system and had established inter-coder percent agreement figures of 80% or more in each of the major categories before collecting data with the system. These data are presented in Tables 6, 7, and 8. Table 6 shows data from morning observations, Table 7 from afternoon observations, and Table 8 represents reading group data only.

Insert Tables 6, 7, and 8 About Here

There are seven major categories of variables coded by this system: academic response opportunities, student initiated questions and comments, private child-created contacts, private teacher-afforded contacts, self-related questions, opinion questions, and behavior related contacts. Each of these categories has several variables within it. The data for these variables were split into three sections for separate data analyses: morning, whole class observational data; afternoon, whole class observational data; and reading group (opposed to the whole class) observational data.

The most important consideration in interpreting the stability data for these low inference variables seems to be the frequency of occurrence of the particular coded interaction. If a variable is a "low frequency" variable, then even minor instability will be amplified and will likely result in a lowering of the probability of obtaining a significantly high stability correlation coefficient. If, however, the variable is a "high frequency" variable, then there should be none of this spurious lowering the probability of occurence of a significant Pearson r. It should be noted that this "frequency factor" operates independently from, though sometimes in conjunction with, the number of subjects with valid data, i.e., the N, in sometimes decreasing and sometimes increasing the possibility of obtaining a significant correlation. The frequency factor affects the strength of relationships indirectly, but nevertheless with considerable impact.

The first set of data are public academic response opportunity variables. In order for a dyadic interaction to be coded in this section, the teacher must ask an individual child an academic question, to which the child responds either correctly, incorrectly, by saying "I don't know," or with no response at all. Any subsequent teacher feedback would be coded, also. The variables pertaining to type of selection of respondents (preselection, nonvolunteer,

volunteer, call out) were generally unstable for the morning and afternoon general class data, but stable for the reading group data, confirming our expectations that reading groups are in some ways more structured and stable than general class teaching milieus. It does make sense that teacher behaviors would be more consistent in a small group instruction mode, since the teacher has fewer students to deal with and a more structured environment. He or she has more control in setting up questioning strategies.

The next two sets of variables, "difficulty level of questions" and "quality of children's answers," were generally unstable. There are two possible explanations for this occurrence: 1) In the second and third grades, teachers don't consistently ask process, product, and choice questions in similar proportions across time, and 2) children in these teachers' classrooms don't respond with consistent percentages of correct and incorrect answers from one year to the next. "Difficulty level of question," therefore, is probably unstable due to low variance. And, with changing student populations from one year to the next, any one teacher might have students who differed considerably from the previous year's students, thus accounting for the differential quality of children's answers across years.

The next variable section, "teacher reactions to correct answers," provide some confirmation of what was expected in the way of stable teacher behavior. When a child's answer is correct, teachers tended to consistently praise, give no feedback, and, for the afternoon data at least, give process feedback. These significant correlations occurred only in the general class data, indicating that, in small reading groups, teachers tend to selectively provide feedback on an individual basis, dealing with each child differently, thus exhibiting no truly consistent feedback strategies.

For "teachers' reactions to part-correct answers" the morning data show praise, criticism, process feedback, sustaining feedback (total), and "rephrase or gives clue" as being stable. For the afternoon data, "another student calls out" and "rephrases or gives clue" are significantly stable. In reading groups, the variables "gives answer" and "repeats question" are stable. These data suggest that when the child's answer is partly correct, teachers tend to react with consistent strategies much more in the mornings than in the afternoons and in reading groups; frequency of occurrence is probably important here, with more of the "discussion-type" of interactions taking place in the mornings. (For the morning data, five of the 11 variables were stable, while only two of the 11 were stable for both the afternoon and the reading group data.)

For the "teacher reactions to wrong answers" set of variables, criticism, sustaining feedback (total), and "rephrase or gives clue" are stable teacher feedback strategies across years for the morning data set. None of the stability coefficients for the afternoon data in this group was statistically significant. In reading groups, when the child's answer is incorrect, "calls on someone else" and "asks new question" were significantly stable across time. The general rule for these variables seems to be that the frequently occurring types of feedback are stable while the types of teacher responses to wrong answers that occur infrequently tend to be unstable as measured across the two years of the study. Here again, stability differences by time of day probably reflect a higher interaction rate in the morning than in the afternoon.

The "teacher reactions to 'I don't know' or 'No response'" variables show little stability, with only the "calls on someone else" and sustaining feedback (total) variables reaching statistical significance for the afternoon

observational data. Within reading groups, only "repeats questions" was significantly stable across time. These data, both for the morning and the afternoon whole class and reading group data subsets, seem to suggest a genuine inconsistency in the teacher's behaviors when dealing with a child who says "I don't know" or when the student makes no response at all.

The next group of variables concerns the different types of teacher feedback combined across all of the response opportunities. Teacher praise (the most frequent teacher reaction coded) was stable for the morning and afternoon general class data. Process feedback was stable for the afternoon observational data, along with "new questions," "repeat, rephrase, or new question after failure to answer," and "another student calls out." The variable "calls on another student after incorrect answer" was significantly stable for both the morning general class data and the reading group data. Also, in reading groups, "repeat, rephrase, or new question after failure to answer" correlated significantly across the two years of the study.

The last variable in this section, "total number of response opportunities per total teaching time," was highly stable in all three of the data subsets -- morning general class, afternoon general class, and reading group. This indicates that teachers do maintain consistent questioning rates from one year to the next, even though the data for more specific measures show that their particular questioning strategies are not always constant across time.

The next two subsets of variables, student initiated questions and student initiated comments, generally lack stability across the two years of the study. There are two primary reasons for this: 1) These are completely pupil controlled behaviors, i.e., the pupils must initiate either a question or comment before a code can be made in this section; thus with different groups of children being

coded in the two years of data collection, there was no reason to expect stability in these categories of variables; and 2) In the second and third grades, the frequencies of occurrence of these types of interactions is generally low, so that any change from one year to the next would show up in the data as a very weak correlational relationship. This was indeed the case for most of the variables in these sections.

Variables from the "self-related" and "opinion" question categories were unstable except for opinion questions followed by praise -- significantly stable for the morning data only -- and proportion of opinion questions of the (total) academically related questions (significant in the afternoon data subset only). Within the "private dyadic contacts" section, two variables, "proportion of private work contacts to all academic interactions" and "proportion of total procedural contacts to public academic contacts," were highly stable across time in all three behavioral subsets: morning, afternoon, and reading group. This indicates that the teachers were consistent in the apportionment of their time for private work contacts and procedural type interactions. These variables seem to be good measures of some aspects of consistent "teacher style," since they reflect the amount of time spent in private versus public teaching and in procedural contacts versus public academic interactions.

Another stable variable in this subset is the one measuring teacher initiated work contacts with lengthy feedback to the students. This was significantly stable for both the morning and afternoon whole class data, but not for the reading group data. This suggests that teachers who do give extended detailed feedback to students do so with regularity across time. The fact that this variable lacks stability for the reading group data does

make sense, since reading groups afford less opportunity for long, private teacher-student interactions. Since the frequency of occurrence here is low, the stability is much more likely to be low.

Behavior management variables which were significantly stable in one or more of the three observational subsets include: proportion of total behavioral contacts which were praise contacts, percent discipline contacts involving one or more errors in the teacher's judgment, percent of teacher errors which were overreactions (for reading group data only), and percent of total control contacts which were non-verbal. This indicates that teachers who praise good behavior were consistent, that those who made errors in their disciplining did so with regularity, and that non-verbal control was used by some teachers consistently across time.

"Proportion of feedback which involved rephrasing" was stable for the reading gruop data but not for the morning or afternoon whole class data. "Brief feedback divided by total brief and long feedback" was stable for both whole class data sets. "Brief feedback" was a high frequency variable in whole class situations, thus increasing the potential for a significant correlation. Lastly, the proportion of math contacts which were public was stable even though the valid N was low (11), indicating that teachers who were coded for these behaviors in both years of the study (public and private math teaching) tended to be consistent in the percentage of contacts concerning math which were public versus those which were private.

DISCUSSION

In summary, while the percentage of the total number of variables which were significantly stable was fairly low, the percentage of "good" variables which were stable was much higher. By "good" variables, we mean variables

which evidenced several attributes: they are measures of behaviors occurring with relatively high frequencies; they are variables which were measured with high inter-coder reliability; no physical or logistical problems were encountered in the actual coding of these variables; and, they are measures of behaviors or interactions which are primarily "teacher controlled" (as opposed to student controlled -- student populations vary from year to year). Most of the variables which are described by these attributes are stable across time, at least within our sample of teachers who were selected for their consistency in producing student gains.

With the exception of the time utilization and lesson presentation data subsets, the high inference data are generally more stable than the low inference, objective coding measures. The primary reasons for this include the fact that the low inference data represent dyadic interactions, so that there is some dependence on student behavior (which is expected to be unstable); also, the high inference variables tended to be measures of overall "teaching style" and of the general "classroom mood." Thus, the stability of these measures is not usually dependent on frequency of occurrence of a specific behavior, but rather on the coder's perception of the teacher's methods and style. Since the coders were in the classrooms for several hours (10 the first year and 30 the second year), they usually were able to make accurate estimates of the teacher behaviors of interest. Some amount of coder bias (e.g."halo"effect, logical errors, etc.) probably occurred, even though coders were trained and warned repeatedly against this. However, this would have acted to decrease the consistency and stability of the high inference data because different coders saw each teacher both within and across years.

Other factors which could have beer operative in lowering the number of significantly stable variables include the small number of observations for the first year data (a problem directly related to the amount of funding available; not to any methodological oversight), and also, as we've mentioned, the fact that student populations do change and may vary considerably across years.

The school district in which the study was conducted mandated several curriculum changes in between the first and second year of the study. These changes included more team teaching, more use of learning centers, more division of responsibilities according to the teacher's strengths (in the second year some teachers taught no math while others taught no language arts), and generally more emphasis on "individualization of instruction." This often entailed individual contracts, more small group work, etc. The period of time that the study was conducted, then, actually was a "time for change" for most, if not all, of the teachers in our sample.

The low inference data are represented by three data subsets: morning, whole class; afternoon, whole class; and small group (reading group) data. We know that many teachers do conduct different activities in the mornings and afternoons. Reading groups also often differ structurally from whole class-type discussions. This is evidenced by somewhat different patterns of correlations for the three groups of data (compare Tables 6, 7, and 8). There are, however, many commonalities between these data sets, too. Some of the "good" variables were significantly stable across time in all three subsets. Often many of the stability by subset differences were meaningfully interpretable, though sometimes this was not the case (see the discussion under the "Low Inference" section). The differences by contexts have been systema-

tically examined (Brophy, Evertson, Crawford, King, and Senior, in press) resulting in the conclusion that, while most of the correlations across contexts are in a positive direction, enough of them were insignificant that there must exist some genuine instability according to the different contexts studied.

Reasons for stability (or absence of stability) given thus far have been centered on non-psychological factors, such as low N, low variance, infrequent occurrence of coded behaviors, context differences, etc. One psychological hypothesis concerning which process variables might be stable was that behaviors and methods which teachers were most aware of would be more stable across time than those they were less aware of. However, this has not proven to be the case. Evertson, Brophy, and Crawford (Note 4) report that most of the teacher process data, including many variables which were shown to be stable across years, were found to be unrelated to teacher questionnaire and interview presage data. Thus, although some evidence of teacher awareness (consonance between observed behavior and self-report of such behavior) was observed, there was little support for the hypothesis that awareness would produce stability automatically.

Thus, we see that in collecting and analyzing classroom process data, one must consider both the context and the time span of data collection. This has implications for data collected in different classes or situations with the same teachers, and also for studies conducted over substantial periods of time -- one can assume the stability-across-time effects to be important for a study spanning a six-or nine-month period, and certainly important for any two-year replication study. Our data suggest that these are two (of many) complexities in multivariate observational research which are worth further study.

Reference Notes

1. Brophy, J. & Evertson, C. Process-product correlations in the Texas teacher effectiveness study: Final report (Res. Rep. 74-4). Austin, Texas: Research and Development Center for Teacher Education, 1974.

2. Brophy, J. & Evertson, C. Texas teacher effectiveness study: Classroom coding manual (Res. Rep. 76-2). Austin, Texas: Research and Development Center for Teacher Education, 1973.

3. Brophy, J. & Evertson, C. The Texas teacher effectiveness project: Presentation of non-linear relationships and summary discussion. (Res. Rep. 74-6). Austin, Texas: Research and Development Center for Teacher Education, 1974.

4. Evertson, C., Brophy, J. & Crawford J. The Texas teacher effectiveness project: An investigation of selected presage-process relationships. (Res. Rep. 75-16). Austin, Texas: Research and Development Center for Teacher Education, 1975.

References

Brophy, J. Stability of teacher effectiveness. American Educational Research Journal, 1973, 10, 245-252.

Brophy, J. & Good, T. The Brophy-Good dyadic interaction system. In A. Simon and E. Boyer (eds.). Mirrors for behavior: An anthology of observation instruments continued, 1970 supplement, Volume A. Philadelphia: Research for Better Schools, Inc., 1970.

Brophy, J., Evertson, C., Crawford, J., King, C., & Senior, K. Stability of measures of classroom process behaviors across three different contexts in second and third grade classrooms. JSAS Catalog of Selected Documents in Psychology, in press.

Emmer, E. & Peck, R. Dimensions of classroom behavior. Journal of Educational Psychology, 1973, 64, 223-240.

Evertson, C. & Brophy, J. High-inference behavioral ratings as correlates of teaching effectiveness. JSAS Catalog of Selected Documents in Psychology, 1973, 3, 97 (Ms. No. 435).

Sherman, G., Brophy, J., Crawford, J. and Evertson, C. Traditional attitudes and teacher consistency in producing student learning gains in the early elementary grades. Journal of School Psychology, in press.

Soar, R. Teacher behavior related to pupil growth. International Review of Education, 1972, 18, 508-526.

Veldman, D. & Brophy, J. Measuring teacher effects on pupil achievement. Journal of Educational Psychology, 1974, 66, 319-324.

Footnotes

The authors wish to acknowledge and thank the following individuals who made special contributions to the research described herein and/or to the preparation of this report:

Carol King, Nancy Moore, Gwen Newman, Brian Peck, Dr. Teresa Peck, Dr. Donald Veldman, and Sidney Weaver.

Special thanks go to Cynthia Coulter, Linda Harris, and Gael Sherman for their help in table preparation.

Table 1
STABILITY CORRELATIONS AND DISTRIBUTION DATA
FOR CLASSROOM OBSERVATION SCALES[1]

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR I RATER AGREEMENT[2]	YEAR II N	YEAR II N-MODE N	YEAR II MODE	YEAR II RATER AGREEMENT[2]
1.	High level of student attention	71	19	.00			NONE	88			NONE	89
2.	Teacher often addresses questions or problems to the whole class	59	19	.01			NONE	77			NONE	73
3.	Teacher is task oriented, doesn't waste time	88	19	.00	19	16	4.1	83	19	16	4.4*	82
4.	Frequent pupil-to-pupil interaction (class relevant)	04	19	.85	19	16	4.4	79			NONE*	87
5.	% of time teacher lectures or demonstrates	15	19	.54			NONE	89	19	15	1.8	62
6.	Negative affect: criticism, hostility	66	19	.00	19	16	4.5	83	19	16	2.0	88
7.	Positive affect: praise, support	68	19	.00	19	15	3.7	86	19	15	2.8	81
8.	Requires high level of generalization, inference, or explanation	64	19	.00			NONE	83			NONE	12
9.	High student withdrawal, or repetitive behavior	14	19	.59	19	16	4.3	75	19	15	1.6	69
10.	Clarity: students show clear understanding of teacher presentations	64	19	.00	19	15	4.9	86	19	14	4.5	60
11.	Enthusiasm: teacher shows enthusiasm, excitement, enjoyment	58	19	.01			NONE	95	19	16	2.8	68
12.	Convergent questioning: most questions have clear-cut correct answers	55	19	.01			NONE	72			NONE	48

Table 1 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	N-MODE N	MODE	RATER AGREEMENT[2]	YEAR II N	N-MODE N	MODE	RATER AGREEMENT[2]

[1]High inference five-point scale ratings made several times during each observation.

[2]First year rater agreement figures are percentage agreement figures for agreement within one point between two raters; second year data are intra-class correlations of ratings of videotapes using all the raters' data.

* Significantly skewed: skewness $p < .05$

NONE No one mode: Bi- or multimodal

Table 2
STABILITY CORRELATIONS AND DISTRIBUTION DATA
FOR TEACHER BEHAVIOR CHECKLIST VARIABLES[1]

NO.	VARIABLE NAME	r	N	p	YEAR I N	N-MODE N	MODE	RATER[2] AGREEMENT Both	One	None	YEAR II N	N-MODE N	MODE	RATER[2] AGREEMENT Both	One	None
A.	Methods of Handling Catch-Up Work															
1.	No remediation; child skips missed work	00	19	1.00			NV	0	2	108	19	7	0	1	14	92
2.	Child must make up work but is not given help	13	19	.60	19	5	0*	1	5	104	19	12	0	3	22	80
3.	Teacher explains work and has child do part of it	26	19	.28	19	8	0*	1	13	96	19	13	4	19	26	44
4.	Another child is assigned to help	-12	19	.64	19	3	0*	0	6	104	19	11	0	3	20	82
5.	Child put in slower group temporarily	00	19	1.00			NV	1	0	109			NV	0	0	108
6.	Other	00	19	1.00			NV	0	1	109	19	3	0*	0	6	102
B.	Rules Regarding Physical Movement															
7.	Must always get permission to leave seat	69	19	.00	19	5	0*	2	10	97	19	5	0*	0	8	100
8.	One at a time without permission	00	19	1.00			NV	0	1	108	19	4	0*	0	15	93
9.	As many as 4 or 5 without permission	13	19	.60	19	3	0*	0	5	104	19	5	0*	0	6	102
10.	Can go quietly to specified places without permission at any time	19	19	.44	19	12	3	23	20	66	19	11	4	26	23	33
11.	No restrictions	58	19	.01	19	8	0*	4	12	93			NV	4	3	97
12.	Some children allowed free movement but not not others	92	19	.00			NV	1	1	107	19	4	0*	2	3	101

Table 2 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR I RATER[2] AGREEMENT Both	YEAR I RATER[2] AGREEMENT One	YEAR I RATER[2] AGREEMENT None	YEAR II N	YEAR II N-MODE	YEAR II MODE	YEAR II RATER[2] AGREEMENT Both	YEAR II RATER[2] AGREEMENT One	YEAR II RATER[2] AGREEMENT None
13.	Only monitors allowed free movement	00	19	1.00			NV	0	1	108			NV	0	0	108
14.	Other	-10	19	.68	19	3	0*	1	3	105			NV	0	4	104
C.	Punishments Used By Teacher															
15.	Stay after school	-19	19	.45	19	6	0*	5	8	96	19	12	0	5	17	81
16.	Spanking	20	19	.41			NV	1	3	105	19	5	0*	4	7	93
17.	Writing sentences on board	00	19	1.00			NV	0	0	109			NV	2	2	102
18.	Isolation within the classroom	37	19	.12	19	9	0	4	14	91			NONE*	6	12	84
19.	Removal from the classroom	19	19	.44	19	6	0*	9	7	93	19	10	0	7	16	78
20.	Note to parents	00	19	1.00			NV	0	1	108	19	9	0*	3	9	93
21.	Send to principal	-19	19	.43	19	3	0*	3	4	102	19	5	0*	1	8	98
22.	Extra reading, math, etc., work	-08	19	.74			NV	0	5	104			NV	0	3	105
23.	Peer pressure (e.g. "You lost the race for your group.")	39	19	.09	19	8	0	1	15	93	19	8	0*	2	19	85
24.	Scolding	55	19	.02			NONE	12	23	74	19	9	4	26	18	38
25.	Discussion of Incident (No Scolding)	44	19	.06	19	8	0*	8	14	87			NONE	17	31	43
26.	Other	07	19	.78	19	9	0*	1	13	95	19	6	0*	0	12	96

Table 2 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR I RATER[2] AGREEMENT Both	YEAR I RATER[2] AGREEMENT One	YEAR I RATER[2] AGREEMENT None	YEAR II N	YEAR II N-MODE	YEAR II MODE	YEAR II RATER[2] AGREEMENT Both	YEAR II RATER[2] AGREEMENT One	YEAR II RATER[2] AGREEMENT None
D.	Rewards Used by Teacher															
27.	Classmates clap or cheer	55	19	.02	19	9	0*	3	16	90	19	9	0*	6	10	86
28.	Special privileges	-10	19	.68	19	4	0*	0	7	102	19	9	2	5	25	73
29.	Waiver or reduction of assignments	-16	19	.51			NV	0	1	108	19	7	0*	0	10	98
30.	Symbols (stars, smiling faces, etc.)	32	19	.18	19	7	0*	3	13	93	19	11	0	3	15	87
31.	Token redeemable for other rewards	00	19	1.00			NV	2	1	106			NV	0	0	108
32.	Concrete rewards (candy, money, prizes)	-11	19	.67			NV	0	3	106	19	4	0*	3	7	95
33.	Jobs (monitor, helper, clean erasers)	-01	19	.96	19	6	0*	2	5	102	19	12	0	3	19	83
34.	Public recognition (gets to read or work problem on board	-02	19	.93	19	12	0	8	16	85	19	12	2	12	29	55
35.	Other	-13	19	.60	19	5	0	0	18	101			NV	1	6	100
E.	Appropriateness of Assignments															
36.	Too short or easy	56	19	.01	19	12	0	13	18	17	19	9	2	8	31	61
37.	Boring, repetitive, monotonous	-18	19	.46	19	13	0	7	20	81	19	12	2	7	24	70
38.	Too hard: students can't get started or continually need help	35	19	.14	19	11	0	3	17	88	19	10	0	10	15	73

Table 2 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR I RATER[2] AGREEMENT Both	YEAR I RATER[2] AGREEMENT One	YEAR I RATER[2] AGREEMENT None	YEAR II N	YEAR II N-MODE	YEAR II MODE	YEAR II RATER[2] AGREEMENT Both	YEAR II RATER[2] AGREEMENT One	YEAR II RATER[2] AGREEMENT None
39.	Continues activity too long, until gets boring	01	19	.98	19	8	0*	5	18	86	19	13	2	6	22	74
40.	No inappropriate assignments	61	19	.01	19	11	0	11	17	80	19	12	0	8	27	65
F.	Distractions: What Do Students Do When Not Working															
41.	Use washroom	20	19	.41			NONE	14	18	76	19	13	0	17	18	56
42.	Repeatedly get supplies for free time activities	13	19	.61	19	4	0*	6	10	92			NONE	8	23	69
43.	Watch reading group or other activity	-24	19	.24	19	9	0	3	20	85			NONE	5	26	72
44.	Talk	26	19	.29	19	8	4	33	20	55	19	4	4*	47	5	9
45.	Play	55	19	.01	19	12	0	12	20	76	19	13	4	20	20	48
46.	Daydream	03	19	.90	19	10	0	11	19	78	19	12	0	13	23	59
47.	Ask for help or look more closely at work on board	-22	19	.38	19	9	0	6	18	84			NONE	6	23	73
48.	Disrupt other students	24	19	.33	19	13	1	15	19	74	19	11	4	26	18	38
49.	Distractions; other	64	19	.00	19	3	0*	0	7	101			NV	0	3	105

Table 2 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR I RATER[2] AGREEMENT Both	YEAR I RATER[2] AGREEMENT One	YEAR I RATER[2] AGREEMENT None	YEAR II N	YEAR II N-MODE	YEAR II MODE	YEAR II RATER[2] AGREEMENT Both	YEAR II RATER[2] AGREEMENT One	YEAR II RATER[2] AGREEMENT None
G.	Student Attitudes															
50.	When having trouble students concentrate or seek help	01	19	.97	19	7	2	13	30	62	19	10	3	18	32	40
51.	When having trouble students merely copy from neighbor	37	19	.12	19	6	0*	8	16	81			NONE	4	14	86
52.	Students work as well when not watched as when watched	18	19	.47			NONE	11	16	78	19	11	2	7	32	62
53.	Students "act up" when unwatched	48	19	.01	19	10	0	11	18	76	19	13	4	15	19	59
54.	Students seem amused by teacher	48	19	.04	19	4	0*	4	8	93	19	3	0*	3	8	94
55.	Students seem to fear teacher	30	19	.21	19	3	0*	0	5	100	19	3	0*	2	5	99
56.	Students seem to respect teacher	71	19	.00	19	9	4	28	12	65	19	9	4	36	12	24
H.	Free Time Materials Available (Not Necessarily Used)															
57.	Books	33	19	.17	19	7	4	41	16	45	19	5	4*	46	8	8
58.	Learning centers (any)	13	19	.60	19	11	0	9	21	72	19	13	2	11	21	65
59.	Listening centers	19	19	.44	19	13	4	19	14	69	19	13	3	20	22	46

Table 2 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	N-MODE	MODE	RATER[2] AGREEMENT Both	One	None	YEAR II N	N-MODE	MODE	RATER[2] AGREEMENT Both	One	None
60.	Visual (picture files, filmstrips)	36	19	.13	19	12	1	4	19	79	19	10	2	10	25	63
61.	Science demonstrations or experiments	-32	19	.18	19	11	0	5	16	81	19	10	0	1	20	86
62.	Other (specify)	24	19	.32	19	5	0*	1	9	92	19	7	0	0	12	96
63.	Coloring pictures	-07	19	.78	19	9	1	7	27	68	19	12	3	19	28	42
64.	Painting, art activities	06	19	.81	19	10	0	8	19	75	19	11	2	9	27	63
65.	Games (any)	-09	19	.71			NONE	14	22	66	19	13	2	12	24	60
66.	Instructional games	06	19	.79			NONE	23	18	61			NONE	26	22	34
67.	Non-instructional games	12	19	.63			NONE	14	24	64	19	11	2	17	25	49
68.	Aquariums, other "looking" exhibits	42	19	.07	19	13	0	18	16	68	19	8	0*	10	8	80
I.	Free Time Materials Observed In Use															
69.	Books	20	19	.41	19	11	4	25	16	61	19	11	4	30	21	27
70.	Learning centers (any)	-13	19	.59	19	7	0*	4	15	83	19	12	2	9	22	68
71.	Listening centers	05	19	.84	19	11	0	9	17	76	19	13	2	14	21	59

Table 2 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR I RATER[2] AGREEMENT Both	YEAR I RATER[2] AGREEMENT One	YEAR I RATER[2] AGREEMENT None	YEAR II N	YEAR II N-MODE	YEAR II MODE	YEAR II RATER[2] AGREEMENT Both	YEAR II RATER[2] AGREEMENT One	YEAR II RATER[2] AGREEMENT None
72.	Visual (picture files, filmstrips)	11	19	.65	19	11	0	3	14	85	19	12	2	7	21	73
73.	Science demonstrations or experiments	21	19	.40	19	8	0	0	15	87	19	5	0*	0	12	96
74.	Other (specify)	30	19	.20	19	6	0*	1	10	91	19	6	0*	0	9	99
75.	Coloring pictures	-06	19	.79			NONE	5	23	74	19	12	2	18	28	44
76.	Painting, art activities	-08	19	.75	19	6	0*	5	15	82			NONE	8	23	69
77.	Games (any)	13	19	.61	19	11	0	6	23	73	19	13	1	8	25	67
78.	..structional games	36	19	.13	19	11	0	12	16	74	19	13	2	19	27	43
79.	Non-instructional games	23	19	.34	19	12	0	12	19	71	19	10	2	14	26	54
80.	Aquariums, other "looking" exhibits	04	19	.86	19	11	0	4	13	85	19	5	0*	5	5	93
J.	Use of Peer Tutoring	67	19	.00	19	9	0	7	15	87	19	8	4	5	14	84
K.	Assigns Homework Besides Seatwork	45	19	.05	19	12	0	4	23	82			NONE	8	29	63
L.	Teacher Sometimes Underreacts to Control Problems, So Serious Problems Go Unresolved	80	19	.00	19	5	0*	6	13	90	19	8	4*	3	9	93

Table 2 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	N-MODE N	MODE	RATER[2] AGREEMENT Both	One	None	YEAR II N	N-MODE N	MODE	RATER[2] AGREEMENT Both	One	None

1
High inference ratings made on checklists by two observers who had each observed the teacher, averaged across observers.

2
Rater agreement data reflect category usage by observers for this subset. The "Both" column indicates the number of times that both raters checked the category in rating a teacher; the "One" column indicates the number of times that one rater coded the category and the other rater did not; the "None" column shows how often neither rater used the category.

NV
No meaningful variance (only 0, 1, or 2 non-modal scores)

*
Significantly skewed: skewness $p < .05$

NONE
No one mode: Bi- or multimodal

Table 3
STABILITY CORRELATIONS AND DISTRIBUTION DATA
FOR HIGH-INFERENCE RATINGS[1]

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR I RATER[2] AGREEMENT	YEAR II N	YEAR II N-MODE	YEAR II MODE	YEAR II RATER[2] AGREEMENT
1.	Typical affectionateness level	68	17	.00	17	13	9.0*	86	19	15	9.6*	80
2.	Most intense affection expression observed	21	17	.42			NONE	90	19	14	10.5*	89
3.	Most intense negative affect observed	04	17	.87			NONE	75	19	16	4.0	76
4.	Solidarity with class: teacher identifies, promotes "we" feeling	-11	17	.66	17	12	4.0	84	19	14	4.3	78
5.	Patient and supportive when correcting	-25	17	.30	17	12	5.0	86	19	15	4.5	78
6.	Students allowed choice in assignments	05	17	.84			NONE	82	19	14	2.0	79
7.	Accepts student ideas and/or integrates them into discussion	-13	16	.64	16	10	4.0	79	19	14	4.0*	92
8.	Admits own mistakes; laughs at self or uses occasion to teach or motivate	-01	16	.97	16	8	4.0	84	19	14	3.5	74
9.	Usually bends close, gets down to child's level	03	17	.89	17	12	4.0	71	19	13	4.5	76
10.	Goes to seats to check work; doesn't stay at desk	15	17	.58	17	14	4.5	83	19	15	3.8	78
11.	Usually speaks to individuals rather than whole class	-25	17	334	17	11	1.5*	84	19	12	2.5	67
12.	Uses advance organizers in introducing activities	28	17	.28	17	13	5.0	71	19	15	2.8	70
13.	Gives complete, detailed instructions: prevents errors before they happen	-12	17	.64	17	11	4.0	67			NONE	82

Table 3 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR I RATER AGREEMENT	YEAR II N	YEAR II N-MODE	YEAR II MODE	YEAR II RATER AGREEMENT
14.	Students eager to respond; no fear	44	17	.07	17	11	4.0	86	19	15	4.0	76
15.	Teacher waits patiently if student doesn't respond promptly	16	16	.57			NONE	73	19	14	4.3	76
16.	Non-competitive atmosphere; no signs of eagerness to see others fail	24	17	.37	17	10	3.0	50	19	12	2.8	70
17.	Students allowed to work in cooperative groups	01	17	.96	17	11	1.0	71	19	15	2.0	56
18.	Teacher recognizes good thinking even when it doesn't lead to "right" answers	14	17	.60			NONE	83	19	15	3.8	65
19.	Democratic leadership style: students share in planning and decision making	13	15	.65	15	10	1.0	81	19	16	3.5	59
20.	Few restrictions on students during seatwork periods	53	15	.04	15	11	3.5	69	19	14	4.0	69
21.	Students expected to care for needs without getting permission	23	17	.37	17	11	5.0	66			NONE	66
22.	Teacher concerned with substantive content, not form, of student responses	26	17	.31	17	11	3.0	75	19	15	3.5	73
23.	Teacher stresses factual realism, rejects or corrects childish idealism	54	15	.04	15	10	3.0	86	19	14	3.0	80
24.	Teacher credibility: students seem to believe and respect teacher	82	17	.00	17	9	5.0*	76	19	8	5.0*	87

Table 3 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR I RATER AGREEME T	YEAR II N	YEAR II N-MODE N	YEAR II MODE	YEAR II RATER AGREEMENT
25.	Showmanship: teacher is melodramatic, expressive, gushy, emotive	60	17	.01	17	11	1.0	80	19	14	1.5*	70
26.	Teacher gets attention before starting, doesn't try to talk over din	70	17	.00			NONE	82	19	15	4.0	76
27.	Chaotic, unplanned, poorly scheduled	-01	17	.98	17	8	3.0	78	19	12	3.0	89
28.	Teacher seems confident, self-assured	79	17	.00	17	7	5.0*	90	19	7	5.0*	93
29.	Politeness: teacher regularly says "please" "thank you," etc.	28	17	.28	17	12	3.5	98	19	15	3.8	76
30.	High concern about achievement	56	17	.02	17	13	4.0	88	19	15	4.0	89
31.	Room is attractive	52	17	.03			NONE	74	19	14	3.5	92
32.	Teacher gives much encouragement to students	48	17	.05	17	12	3.0	88	19	14	3.8	91
33.	Room is uncrowded	24	17	.35	17	10	3.0	74	19	15	2.5	73
34.	Teacher explains rules or decisions when reasons aren't obvious	32	16	.22	16	10	5.0*	87			NONE	70
35.	Teacher well organized, prepared	18	17	.50	17	9	5.0	84			NONE	80
36.	Teacher regularly monitors class, knows what's going on	79	17	.00	17	12	4.5	86	19	14	4.5	90
37.	Smooth, efficient transitions, little time wasted	43	17	.08			NONE	70			NONE	65

Table 3 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR I RATER AGREEMENT	YEAR II N	YEAR II N-MODE N	YEAR II MODE	YEAR II RATER AGREEMENT
38.	Monitors determined "automatically" by a systematic procedure	45	17	.07	17	9	5.0*	74	19	15	4.8	85
39.	"Busy", cluttered classroom	45	17	.07	17	10	3.0	86	19	14	3.8	73
40.	Students compliant, obedient	68	17	.00	17	12	5.0	94	19	13	5.0	94
41.	Teacher gives overly explicit, repetitious directions	-24	17	.36	17	12	3.0	74			NONE	76
42.	Well established routines minimize interruptions; room runs "automatically"	69	17	.00	17	13	4.0	82			NONE	87

[1] High inference ratings on 13 point scales (first 3 variables) or 5-point scales (all other variables) made by two observers who had each observed the teacher, averaged across coders.

[2] Agreement data reflect percent agreement within 2 points on the first 3 variables and within 1 point on all other variables (using the two raters' data who had observed the same teacher).

* Significantly skewed: skewness $p < .05$

NONE No one mode: Bi- or multi-modal

Table 4
STABILITY CORRELATIONS AND DISTRIBUTION DATA
FOR TIME UTILIZATION VARIABLES[1]

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II MODE
1.	Total available time	-04	19	.85	19	14	697.8	19	15	2197.8*
2.	Total teacher controlled time	16	19	.53			NONE	19	16	1992.0
3.	Proportion of total time structured by teacher	02	19	.93			NONE			NONE
4.	Proportion of structured time in language arts	-59	16	.02	19	17	.06			NONE
5.	Proportion of structured time in math	37	16	.16	19	16	.19			NONE
6.	Proportion of structured time in art	64	16	.01	19	7	0*			NONE
7.	Proportion of structured time in spelling	46	16	.07	19	16	.11	16	12	.07
8.	Proportion of structured time in reading groups	27	16	.31	19	15	.27			NONE
9.	Proportion of structured time in social studies	18	16	.50	19	12	0*			NONE
10.	Proportion of structured time in transitions	63	16	.01			NONE	16	12	.03
11.	Proportion of structured time in morning routine	-43	16	.09	19	12	0*	16	12	.01
12.	Proportion of structured time in special activities.	05	16	.86	19	16	.31	16	12	.06

[1] The first two variables reflect "number of minutes" -- taken from coding sheets filled out in the classrooms. The rest of the variables are proportions computed by dividing the number of minutes spent in a specific activity by the total amount of structured time.

* significantly skewed: skewness $p < .05$

NONE No one mode: Bi-or multi-modal

Table 5
STABILITY CORRELATIONS AND DISTRIBUTION DATA
FOR PERCENTAGE ESTIMATE VARIABLES[1]

NO.	VARIABLE NAME	r	N	p	YEAR I N	N-MODE N	MODE	RATER[2] AGREEMENT Both	One	None	YEAR II N	N-MODE N	MODE	RATER[2] AGREEMENT Both	One	None
A.	Methods Used to Call for Attention															
1.	Says nothing, waits for quiet	39	19	.10			NONE	21	16	15	19	15	0	32	8	36
2.	Raps desk lightly, uses normal voice	37	19	.11	19	7	0*	1	16	35	19	9	0*	2	11	93
3.	Gimmick (light flick, bell, clicker)	27	19	.27	19	6	0*	5	8	39	19	6	0*	7	8	86
4.	Raises voice over the din	19	19	.44			NONE	22	22	8			NONE	25	11	47
5.	Raises voice and singles out individuals	12	19	.64	19	15	0	12	23	17			NONE	33	11	31
6.	Shouts, becomes angry, or scolds class	57	19	.01	19	4	0*	5	5	42	19	9	0	14	4	76
7.	Shouts, becomes angry, or scolds individuals	61	19	.01	19	5	0*	5	6	41	19	10	0*	14	9	71
8.	Whispers or speaks softly to nearby pupils (at first)	-13	19	.61			NV	0	5	47	19	6	0*	5	5	93
9.	Other (includes any method not listed above)	58	19	.01	19	6	0*	3	12	37	19	5	0*	1	6	100
B.	% Attending															
10.	Estimated % of students paying attention	64	19	.00	19	16	83.0	50	0	0	19	15	90.7*	54	0	0

Table 5 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	N-MODE	N	MODE	RATER[2] AGREEMENT Both	One	None	YEAR II N	N-MODE	N	MODE	RATER[2] AGREEMENT Both	One	None
C.	What Does the Teacher Do When a Child Doesn't Understand?																	
11.	Stops what she's doing, explains	16	19	.52				NONE	29	11	1	19	16		53.0	42	12	12
12.	Delays child then explains later	62	19	.01	19	16		13.0	21	16	4	19	15		20.1	31	18	28
13.	Delays, but then fails to follow up	76	19	.00	19	7		0*	3	10	28	19	8		0*	10	7	81
14.	Asks another child to explain	0	19	.99	19	15		0	8	22	10	19	16		2.8	14	25	55
15.	Scolds child for not understanding	-11	19	.66	19	5		0*	3	10	27	19	10		0*	4	16	84
16.	Encourages child but gives no help	-24	19	.33	19	8		0*	3	14	23	19	14		0	8	22	70
17.	Refuses help ("You're on your own.")	20	19	.43	19	6		0*	2	13	25	19	8		0*	3	11	91
18.	Sends child to aid or other adult	18	19	.48				NV	0	2	38	19	5		0*	0	10	98
19.	Other (Includes any method not listed above)	0	19	1.00				NV	0	4	36				NV	0	1	107
D.	Locations of Interactions																	
20.	Child goes to teacher's desk	-22	19	.37				NONE	50	0	0	19	16		37.2	52	2	2
21.	Teacher goes to child's desk	-21	19	.38				NONE	50	0	0	19	16		62.8	52	2	2

Table 5 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR I RATER[2] AGREEMENT Both	YEAR I RATER[2] AGREEMENT One	YEAR I RATER[2] AGREEMENT None	YEAR II N	YEAR II N-MODE	YEAR II MODE	YEAR II RATER[2] AGREEMENT Both	YEAR II RATER[2] AGREEMENT One	YEAR II RATER[2] AGREEMENT None
E.	What Teacher Does When Child Is Stuck While Reading In Reading Group															
22.	Gives word	12	18	.64	18	15	50.0	26	3	0	19	16	44.0	42	12	12
23.	Gives first sound or syllable	35	18	.15	18	15	7.0	12	9	8	19	16	15.6	18	25	47
24.	Child starts sentence or paragraph over	11	18	.67	18	9	0*	0	12	17	19	9	0*	2	15	89
25.	Gives context clue or definition	-19	18	.46	18	8	0*	4	5	20	19	15	0	13	32	50
26.	Asks another child to give word	03	18	.91	18	13	0	15	7	7	19	15	11.6	25	22	36
27.	Gives clue unrelated to sound or meaning ("It's one of our new words.")	20	18	.42	18	4	0*	1	7	21	19	12	0	6	20	76
28.	Tells child to skip, go to next word	-08	18	.75			NV	0	2	27			NV	0	4	104

[1] Two observers who each had observed the same teacher estimated percentages for the various categories; scores were then obtained by averaging.

[2] The rater agreement data reflect category usage by the observers. The "Both" column indicates the number of times that both raters gave a non-zero rating; the "One" column shows the number of times one coder gave a non-zero rating; the "None" column indicates the number of times neither observer gave a non-zero rating of the teacher for that variable.

NV No meaningful variance (only 0, 1, or 2 non-modal scores)

* Significantly skewed: skewness $p < .05$

NONE No one mode: Bi- or multimodal

Table 6
STABILITY CORRELATIONS AND DISTRIBUTION DATA
FOR TEACHER PROCESS VARIABLES FROM THE
EXPANDED BROPHY-GOOD DYADIC OBSERVATIONAL SYSTEM: MORNING OBSERVATIONS [1]

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II MODE
A.	Selecting Respondents to Questions									
1.	% Preselects respondent before asking questions	36	18	.14	18	11	0*	19	14	0*
2.	Calls on non-volunteer	24	18	.34			NONE	19	16	.24
3.	Calls on volunteer	13	18	.62	18	15	.43			NONE
4.	Student calls out answers	07	18	.79			NONE			NONE*
B.	Difficulty Level of Questions									
5.	Process questions/process + product questions	12	18	.64	18	12	0	19	15	.02*
6.	Choice questions/process + product + choice	02	18	.93	18	15	.04	19	16	.17*
C.	Quality of Children's Answers									
7.	% Correct	31	17	.22			NONE	19	16	.69
8.	% part-correct	57	17	.02	18	14	0			NONE
9.	% wrong	11	17	.69	18	15	.11*			NONE
10.	% "don't know"	-02	17	.95	18	7	0*	19	9	0*
11.	% no response	70	17	00	18	14	.07			NONE

Table 6 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
D.	Teacher Reactions to Correct Answers									
12.	Praise	79	18	.00	18	12	0*	19	16	.07*
13.	Criticism	11	18	.67			NV			NV
14.	Failure to give feedback	41	18	.08	18	6	0*	19	3	0*
15.	Process feedback	33	18	.17	18	4	0*	19	10	0*
16.	New question	-11	13	.73	18	14	0	19	9	0*
E.	Teacher Reactions to Part-Correct Answers									
17.	Praise	82	18	.00			NV			NONE*
18.	Criticism	100	12	.00			NV			NV
19.	Failure to give feedback	00	13	1.00			NV			NV
20.	Process feedback	62	13	.02	14	3	0*			NV
21.	Gives the answer	02	13	.94	14	9	0	17	9	0*
22.	Calls on someone else	10	13	.74	14	7	0*			NONE
23.	Another student calls out the answer	-11	13	.72			NV	17	5	0*
24.	Repeats, rephrases, or asks new question	64	13	.02	14	9	0*			NONE
25.	Repeats question	26	13	.39	14	6	0*	17	8	0*

Table 6 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
26.	Rephrases or gives clue	58	13	.04	14	5	0*	.17	13	.01*
27.	Asks new question	-01	13	.98	14	5	0*	17	6	.01*
F.	Teacher Reactions to Wrong Answers									
28.	Praise	00	16	1.00	16	3	0*			NV
29.	Criticism	79	16	.00			NV	19	4	0*
30.	Failure to give feedback	22	16	.43			NV	19	5	0*
31.	Process feedback	-24	16	.36	16	3	0*	19	10	0*
32.	Gives the answer	-33	16	.21	16	6	0*	19	15	0
33.	Calls on someone else	15	16	.59	16	13	.51			NONE
34.	Another student calls out the answer	-16	16	.57			NV	19	8	0*
35.	Repeats, rephrases, or asks new question	50	16	.05			NONE			NONE
36.	Repeats question	06	16	.82			NONE	19	14	0
37.	Rephrases or gives clue	42	16	.09	16	10	0*	19	15	0
38.	Asks new question	-18	16	50	16	7	0*	19	8	0*

Table 6 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
G.	Teacher Reactions to "I Don't Know" or No Response									
39.	Criticism	00	18	1.00			NV			NV
40.	Failure to give feedback	-08	18	.74			NV			NV
41.	Gives the answer	16	18	.54	18	5	0*	19	9	0*
42.	Calls on someone else	32	18	.19	18	14	0	19	15	.49
43.	Another student calls out the answer	-21	18	.40	18	5	0*	19	7	0*
44.	Repeats, rephrases, or asks new question	-10	18	.70	18	12	0*	19	14	0
45.	Repeats question	-02	18	.95	18	11	0	19	11	0
46.	Rephrases or gives clue	-18	18	.48	18	6	0*	19	13	0*
47.	Asks new question	-03	18	.90	18	4	0*	19	8	0*
H.	Teacher Reactions Combined Across All Response Opportunities									
48.	Praise	78	18	.00	18	13	0*	19	15	.06*
49.	Failure to give feedback	-03	18	.90	18	6	0*			NONE*
50.	Process feedback	26	18	.30	18	6	0*	19	14	0*

Table 6 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II MODE
51.	New question	29	18	.24	18	14	0			NONE*
52.	Repeat, rephrase, or new question after failure to answer	28	18	.25	18	15	.29			NONE
53.	Repeats question after failure to answer correctly	-10	18	.71			NONE			NONE
54.	Gives the answer after failure to answer correctly	-05	18	.84	18	11	0*	19	16	.02
55.	Calls on another student after failure to answer correctly	39	19	.10			NONE			NONE*
56.	Another student calls out answer after failure to answer correctly	02	18	.94	18	5	0*	19	12	0*
I.	Student Response Opportunities									
57.	Response opportunities/ total teaching time	.41	19	.08			NONE			NONE
J.	Student Initiated Questions (SIQ's)									
58.	% SIQ's irrelevant	-18	15	.54	15	8	0	18	4	0*
59.	% SIQ's called out	-17	15	.56	15	7	0	18	6	1.00
60.	Praise of question after relevant SIQ	00	15	1.00			NV	18	3	0*

Table 6 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II MODE
61.	Criticism of question after relevant SIQ	-09	15	.75			NV			NV
62.	% relevant SIQ's delayed	-14	15	.61	15	4	0*			NV
63.	% relevant SIQ's not accepted	-03	15	.92	15	3	0*	18	4	0*
64.	% relevant SIQ's given brief feedback	13	15	.64			NONE	18	11	1.00
65.	% relevant SIQ's given long feedback	-17	15	.36	15	10	0	18	9	0
66.	% relevant SIQ's redirected to class	-02	15	.94	15	3	0*	18	3	0*
67.	Behavioral praise of relevant SIQ	00	15	1.00			NV			NV
68.	Behavioral criticism of relevant SIQ	68	15	.01			NV			NV
69.	Behavioral warning after relevant SIQ	00	15	1.00			NV			NV
70.	Criticism of question after irrelevant SIQ	00	02	1.00			NV			NV
71.	% irrelevant SIQ given no feedback	00	02	1.00			NV			NV
72.	% irrelevant SIQ delayed	00	02	1.00			NV			NV
73.	% irrelevant SIQ given brief feedback	00	02	1.00			NONE	4	2	1.00
74.	% irrelevant SIQ given long feedback	00	02	1.00			NV			NV
75.	% irrelevant SIQ not accepted	00	02	1.00	8	3	.02	4	2	0

Table 6 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I N	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II N	YEAR II MODE
76.	% irrelevant SIQ redirected to class	00	02	1.00				NV				NV
77.	Behavioral criticism after irrelevant SIQ	00	02	1.00				NV				NV
78.	Behavioral warning after irrelevant SIQ	00	02	1.00				NV				NV
K.	Student Initiated Public Interactions											
79.	Student initiated comments and questions/total response opportunities	72	19	.00				NONE*	19	15		.05*
L.	Student Initiated Comments (SIC's)											
80.	% SIC's relevant	-14	17	.59	17	11		1.00	19	9		1.00*
81.	% SIC's called out	-46	17	.06	17	10		.01	19	14		1.00
82.	Praise of comment after relevant SIC	-12	17	.64	17	3		0*	19	4		0*
83.	% relevant SIC's given no feedback	17	17	.52	17	3		0*	19	5		0*
84.	% relevant SIC's delayed	-11	17	.69				NV	19	3		0*
85.	% relevant SIC's not accepted	00	17	.98	17	8		0*	19	10		0
86.	% relevant SIC's accepted	14	17	.59	17	10		1.00	19	14		1.00
87.	% relevant SIC's integrated into discussion topic	-33	17	.20	17	7		0*	19	11		0*

Table 6 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I N	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II N	YEAR II MODE
88.	% relevant SIC's which cause a shift in topic	00	17	1.00				NV				NV
89.	Behavioral praise after relevant SIC's	00	17	1.00				NV				NV
90.	Behavioral criticism after relevant SIC's	00	17	1.00				NV				NV
91.	Behavioral warning after relevant SIC's	-09	17	.73				NV				NV
92.	Praise of comment after irrelevant SIC's	00	06	1.00				NV				NV
93.	% irrelevant SIC's given no feedback	00	06	1.00	11	3		0*				NV
94.	% irrelevant SIC's delayed	00	06	1.00				NV				NV
95.	% irrelevant SIC's not accepted	-32	06	.55	11	4		.02*	9	3		0
96.	% irrelevant SIC's accepted	50	06	.55	11	6		0	9	4		1.00
97.	% irrelevant SIC's integrated into discussion topic	-20	06	.70				NV				NV
98.	% irrelevant SIC's which cause a shift in topic	00	06	1.00				NV				NV
99.	Behavioral criticism after irrelevant SIC's	00	06	1.00				NV				NV
100.	Behavioral warning after irrelevant SIC's	00	06	1.00				NV				NV
M.	Self and Opinion Questions											
101.	Self questions/process + product + choice questions	-04	19	.86	19	12		.04*	19	15		.01*

Table 6 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
102.	% self questions which were subject-matter related	00	02	1.00			NV	16	11	1.00
103.	% self questions related to personal preference	00	02	1.00			NV	16	10	0
104.	Opinion questions/process + product + choice questions	-03	18	.92	18	6	0*	19	13	0*
105.	% opinion questions given no response	00	05	1.00			NV			NV
106.	% opinion questions followed by praise	94	05	.02	6	2	0	14	4	0*
107.	% of opinion questions followed by teacher disagreement	00	05	1.00			NV	14	7	0
108.	% student opinions accepted	-01	05	.99	6	4	1.00	14	10	1.00*
109.	% student opinions integrated into discussion topic	00	05	1.00	6	3	0	14	3	0*
N.	Private Dyadic Contacts									
110.	% private contacts student initiated	45	19	.05	19	17	.16			NONE
111.	Student initiated work contacts involving praise	26	19	.28	19	13	0*	19	15	0
112.	Student initiated work contacts involving criticism	36	19	.13	19	8	0*	19	15	0*

Table 6 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I N	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II N	YEAR II MODE
113.	% of private work contacts student initiated	-40	19	.08				NONE				NONE
114.	% student initiated contacts delayed	-11	19	.66	19	14		0*				NONE
115.	% student initiated contacts given brief feedback	12	19	.63				NONE*				NONE
116.	% student initiated contacts given long feedback	-24	19	.32				NONE	19	15		.04*
117.	% student initiated contacts involving personal concerns	-35	19	.14	19	17		.33	19	14		.30*
118.	% student initiated requests delayed	-12	19	.63	19	12		0*				NONE
119.	Private work contacts/private work contacts + public response opportunities	59	19	.01	19	16		.48	19	15		.38
120.	Procedural contacts/procedural contacts + response opportunities	63	19	.00				NONE				NONE
121.	Teacher initiated work contacts/teacher initiated work + procedure contacts	10	19	.69				NONE				NONE
122.	Teacher initiated work contacts involving praise	08	19	.74	19	15		.04*	19	16		.07
123.	Teacher initiated work contacts involving more observation	30	19	.21	19	12		0	19	15		0*

Table 6 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
124.	Teacher initiated work contacts involving brief feedback	-11		.66			NONE			NONE
125.	Teacher initiated work contacts involving long feedback	55	19	.01			NONE*			NONE
126.	% teacher initiated procedural contacts which were management requests	05	19	.82			NONE	19	7	.95*
127.	% teacher thanks student for doing a favor request	36	18	.13	18	9	0*	19	12	0*
128.	% teacher thanks student for following a management request	-27	18	.29	18	4	0*	19	12	0*
O.	Combined Teacher Evaluation Statements									
129.	Academic praise/academic praise + academic criticism	25	19	.31	19	15	1.00			NONE*
130.	Behavioral praise/total behavioral contacts	-06	19	.81	19	6	0*	19	12	0*
131.	Behavioral warnings/behavioral warnings + behavioral criticism	09	19	.71			NONE	19	16	.83
P.	Discipline and Control Errors									
132.	% Discipline contacts involving one or more errors	-13	19	.61	19	12	0*	19	12	0

Table 6 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II MODE
133.	Target errors/total errors	14	09	.72	13	3	0*			NONE
134.	Timing errors/total errors	-08	09	.83	13	9	.01	12	7	0
135.	Overreactions/total errors	-59	09	.09	13	6	0			NONE
136.	Nonverbal control contacts/total control contacts	54	19	.02	19	12	0	19	16	.07*
Q.	Combined Teacher Feedback Data									
137.	Repeat/repeat + rephrase + new question	-12	15	.68			NONE			NONE
138.	Rephrase/repeat + rephrase + new question	-12	15	.68			NONE	18	15	.61
139.	Brief feedback/brief + long feedback	56	19	.01	19	16	.53	19	15	.78
R.	Math Contacts									
140.	Total public math contacts/total public math contacts + total private math contacts	62	11	.04	14	3	1.00*			NONE
141.	Total teacher initiated private math contacts/ total public math contacts + total private math contacts	67	03	.53	4	2	.07	16	14	.28

Table 6 (Continued)

No.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I N	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II N	YEAR II MODE

[1]Process variables for which data were collected in the first and second years of the study came from an expanded version of the Brophy-Good Dyadic Observational System. Variables are either proportions or frequencies from the different interaction categories. These data were taken during morning, whole-class observations.

NV No meaningful variance (only 0, 1, or 2 non-modal scores)

* Significantly skewed: skewness $p < .05$

NONE No one mode: Bi- or Multi-modal

Table 7
STABILITY CORRELATIONS AND DISTRIBUTION DATA
FOR TEACHER PROCESS VARIABLES FROM THE
EXPANDED BROPHY-GOOD DYADIC OBSERVATIONAL SYSTEM: AFTERNOON OBSERVATIONS[1]

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I N	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II N	YEAR II MODE
A.	Selecting Respondents to Questions											
1.	% Preselects respondent before asking questions	-16	17	.60	18	6		0*	18	6		0*
2.	Calls on non-volunteer	44	17	.07				NONE				NONE
3.	Calls on volunteer	00	17	.99	18	15		.50	18	15		0
4.	Student calls out answers	07	18	.79	18	15		0*	18	14		.02*
B.	Difficulty Level of Questions											
5.	Process questions/process + product questions	12	18	.64	18	9		0*	18	13		0
6.	[illegible]ice questions/process + product + choice	02	18	.93	18	12		0*	18	15		.09*
C.	Quality of Children's Answers											
7.	% Correct	31	17	.22	18	15		.76				NONE
8.	% part-correct	57	17	.02	18	15		.07	18	15		0
9.	% wrong	11	17	.69				NONE				NONE
10.	% "don't know"	-02	17	.95	18	8		0*	18	8		0*
11.	% no response	70	17	.00	18	14		.04*				NONE

Table 7 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
D.	Teacher Reactions to Correct Answers									
12.	Praise	79	18	.00	18	15	0	18	13	0*
13.	Criticism	11	18	.67			NV	18	3	0*
14.	Failure to give feedback	41	18	.08	18	7	0*	18	4	0*
15.	Process feedback	33	18	.17	18	4	0*	18	8	0
16.	New question	-11	13	.73	18	15	.05*			NONE*
E.	Teacher Reactions to Part-Correct Answers									
17.	Praise	82	18	.00			NV			NV
18.	Criticism	1.00	12	.00			NV			NV
19.	Failure to give feedback	00	13	1.00	16	3	0*			NV
20.	Process feedback	62	13	.02	16	3	0*	15	5	0*
21.	Gives the answer	02	13	.94	16	6	0*	15	8	0
22.	Calls on someone else	10	13	.74	16	12	0*	15	12	.24*
23.	Another student calls out the answer	-11	13	.72			NV			NV
24.	Repeats, rephrases, or asks new question	64	13	.02	16	12	0			NONE
25.	Repeats question	26	13	.39	16	9	0*	15	7	0*

Table 7 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
26.	Rephrases or gives clue	50	15	.06	16	10	0	15	11	0
27.	Asks new question	16	15	.57	16	7	0	15	8	0*
F.	Teacher Reactions to Wrong Answers									
28.	Praise	00	15	1.00			NV			NV
29.	Criticism	-18	15	.52			NV	16	8	0*
30.	Failure to give feedback	-12	15	.68	17	3	0*			NV
31.	Process feedback	04	15	.88	17	6	0*	16	8	0
32.	Gives the answer	12	15	.68	17	5	0*	16	10	0*
33.	Calls on someone else	15	15	.06	17	13	0	16	13	.58
34.	Another student calls out the answer	-01	15	.97			NV	16	5	0*
35.	Repeats, rephrases, or asks new question	21	15	.46	17	14	0	16	14	.19
36.	Repeats question	-04	15	.87	17	11	0*	16	11	0*
37.	Rephrases or gives clue	06	15	.82	17	12	0			NONE
38.	Asks new question	-06	15	.84	17	9	0*	16	6	0*

Table 7 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II MODE
G.	Teacher Reactions to "I Don't Know" or No Response									
39.	Criticism	-15	15	.60			NV	16	15	0*
40.	Failure to give feedback	00	15	1.00			NV			NV
41.	Gives the answer	10	15	.71	16	4	0*	16	7	0
42.	Calls on someone else	60	15	.02			NONE			NONE
43.	Another student calls out the answer	04	15	.89			NV	16	6	0*
44.	Repeats, rephrases, or asks new question	74	15	.01	16	12	0	16	12	0
45.	Repeats question	-14	15	.62	16	6	0*	16	11	0*
46.	Rephrases or gives clue	-29	15	.30	16	8	0*	16	7	0*
47.	Asks new question	12	15	.68	16	7	0*	16	7	0*
H.	Teacher Reactions Combined Across All Response Opportunities									
48.	Praise	43	17	.08			NONE	18	13	0*
49.	Failure to give feedback	-19	17	.48	18	10	0*	18	6	0*
50.	Process feedback	48	17	.05	18	10	0*	18	13	0

Table 7 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
51.	New question	67	17	.01	18	15	.04			NONE*
52.	Repeat, rephrase, or new question after failure to answer	53	17	.03	17	14	.41			NONE
53.	Repeats question after failure to answer correctly	-18	17	.49	18	12	0*	18	14	0*
54.	Gives the answer after failure to answer correctly	04	17	.86	18	11	0	18	14	0*
55.	Calls on another student after failure to answer correctly	26	17	.32	18	15	.12			NONE
56.	Another student calls out answer after failure to answer correctly	47	17	.06	18	4	0*	18	8	0*
I.	Student Response Opportunities									
57.	Response opportunities/ total teaching time	.57	18	.01	19	16	.47			NONE
J.	Student Initiated Questions (SIQ's)									
58.	% SIQ's irrelevant	-21	13	.50	15	4	0*	15	3	0*
59.	% SIQ's called out	15	13	.64	15	9	1.00	15	11	1.00
60.	Praise of question after relevant SIQ	74	12	.01			NV			NV

Table 7 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II MODE
61.	Criticism of question after relevant SIQ	00	12	1.00			NV			NV
62.	% relevant SIQ's delayed	16	12	.63	14	4	0*			NV
63.	% relevant SIQ's not accepted	17	12	.60			NV	15	7	0*
64.	% relevant SIQ's given brief feedback	-14	12	.67	14	9	1.00	15	11	1.00
65.	% relevant SIQ's given long feedback	-23	12	.47	14	7	0*	15	8	.01*
66.	% relevant SIQ's redirected to class	-13	12	.70			NV	15	4	0*
67.	Behavioral praise of relevant SIQ	00	12	1.00			NV			NV
68.	Behavioral criticism of relevant SIQ	00	12	1.00			NV			NV
69.	Behavioral warning after relevant SIQ	00	12	1.00			NV			NV
70.	Criticism of question after irrelevant SIQ	00	00	1.00			NV			NV
71.	% irrelevant SIQ given no feedback	00	00	1.00			NV			NV
72.	% irrelevant SIQ delayed	00	00	1.00			NV			NV
73.	% irrelevant SIQ given brief feedback	00	00	1.00			NONE			NV
74.	% irrelevant SIQ given long feedback	00	00	1.00			NV			NV
75.	% irrelevant SIQ not accepted	00	00	1.00			NV			NV

Table 7 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II MODE
76.	% irrelevant SIQ redirected to class	00	00	1.00			NV			NV
77.	Behavioral criticism after irrelevant SIQ	00	00	1.00			NV			NV
78.	Behavioral warning after irrelevant SIQ	00	00	1.00			NV			NV
K.	Student Initiated Public Interactions									
79.	Student initiated comments and questions/total response opportunities	09	18	.71			NONE*			NONE*
L.	Student Initiated Comments (SIC's)									
80.	% SIC's relevant	04	17	.86	19	13	1.00*	17	6	1.00
81.	% SIC's called out	23	17	.38	19	13	1.00	17	13	1.00
82.	Praise of comment after relevant SIC	16	16	.57	18	3	0*	17	6	0*
83.	% relevant SIC's given no feedback	-18	16	.52	18	5	0*	17	7	.01*
84.	% relevant SIC's delayed	26	16	.33	18	4	0*	17	6	0*
85.	% relevant SIC's not accepted	-04	16	.88	18	9	0*	17	10	0*
86.	% relevant SIC's accepted	04	16	.87			NONE			NONE
87.	% relevant SIC's integrated into discussion topic	-11	16	.68	18	14	0*	17	12	0

Table 7 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I N	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II N	YEAR II MODE
88.	% relevant SIC's which cause a shift in topic	-08	16	.78				NV				NV
89.	Behavioral praise after relevant SIC's	00	16	1.00				NV				NV
90.	Behavioral criticism after relevant SIC's	-07	15	.80				NV				NV
91.	Behavioral warning after relevant SIC's	00	15	1.00				NV				NV
92.	Praise of comment after irrelevant SIC's	00	07	1.00				NV				NV
93.	% irrelevant SIC's given no feedback	-36	07	.43	13		7	.01*	7		3	0
94.	% irrelevant SIC's delayed	-26	07	.58				NV				NV
95.	% irrelevant SIC's not accepted	00	07	.99	13		5	0				NV
96.	% irrelevant SIC's accepted	-41	07	.36	13		9	.01				NONE
97.	% irrelevant SIC's integrated into discussion topic	20	07	.67	13		4	0*				NV
98.	% irrelevant SIC's which cause a shift in topic	00	07	1.00				NV				NV
99.	Behavioral criticism after irrelevant SIC's	00	07	1.00				NV				NV
100.	Behavioral warning after irrelevant SIC's	00	07	1.00				NV				NV
M.	Self and Opinion Questions											
101.	Self questions/process + product + choice questions	-07	18	.78				NONE*	18		15	0*

Table 7 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II MODE
102.	% self questions which were subject-matter related	00	00	1.00			NV			NONE
103.	% self questions related to personal preference	00	00	1.00			NV	15	11	0
104.	Opinion questions/process + product + choice questions	43	17	.08	18	5	0*	18	11	0*
105.	% opinion questions given no response	00	04	1.00			NV			NV
106.	% opinion questions followed by praise	-42	04	.58			NV	11	8	0*
107.	% of opinion questions followed by teacher disagreement	-36	04	.64			NV	11	4	0*
108.	% student opinions accepted	-78	04	.22			NV	11	7	1.00
109.	% student opinions integrated into discussion topic	-14	04	.85			NV	11	4	0*
N.	Private Dyadic Contacts									
110.	% private contacts student initiated	40	18	.10	19	14	.51	18	15	.32
111.	Student initiated work contacts involving praise	16	18	.52	19	13	0*			NONE
112.	Student initiated work contacts involving criticism	19	18	.47	19	14	0*	18	12	0*

Table 7 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II MODE
113.	% of private work contacts student initiated	-08	18	.74			NONE	18	15	.69
114.	% student initiated contacts delayed	28	18	.25	19	16	.01*			NONE
115.	% student initiated contacts given brief feedback	-17	18	.51			NONE	18	15	.45
116.	% student initiated contacts given long feedback	22	18	.37	19	16	.17			NONE
117.	% student initiated contacts involving personal concerns	35	18	.15			NONE*	18	15	.13
118.	% student initiated requests delayed	06	18	.80	19	9	0*			NONE
119.	Private work contacts/private work contacts + public response opportunities	59	18	.01			NONE	18	15	.47
120.	Procedural contacts/procedural contacts + response opportunities	63	18	.01			NONE	18	15	.39
121.	Teacher initiated work contacts/teacher initiated work + procedure contacts	32	18	.19			NONE			NONE
122.	Teacher initiated work contacts involving praise	01	18	.95	19	15	0*			NONE
123.	Teacher initiated work contacts involving mere observation	-06	18	.80	19	12	.01*	18	13	0*

Table 7 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II MODE
124.	Teacher initiated work contacts involving brief feedback	-15	18	.56			NONE*	18	15	.62*
125.	Teacher initiated work contacts involving long feedback	72	18	.00			NONE*	18	15	.13
126.	% teacher initiated procedural contacts which were management requests	26	18	.31			NONE	18	15	.96*
127.	% teacher thanks student for doing a favor request	39	18	.11	19	7	0*	18	11	0
128.	% teacher thanks student for following a management request	37	18	.13	19	4	0*	18	11	0*
O.	Combined Teacher Evaluation Statements									
129.	Academic praise/academic praise + academic criticism	05	18	.85	19	16	1.00			NONE
130.	Behavioral praise/total behavioral contacts	71	18	.00	19	7	0*	18	10	0*
131.	Behavioral warnings/behavioral warnings + behavioral criticism	18	18	.49	19	16	.50			NONE
P.	Discipline and Control Errors									
132.	% Discipline contacts involving one or more errors	51	18	.03	19	12	0*	18	12	0*

Table 7 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
133.	Target errors/total errors	-25	08	.55	12	3	0*	12	6	0*
134.	Timing errors/total errors	44	08	.28	12	6	.01	12	4	0
135.	Overreactions/total errors	11	08	.80	12	7	0	12	8	1.00
136.	Nonverbal control contacts/total control contacts	-13	18	.62	19	14	0			NONE
Q.	Combined Teacher Feedback Data									
137.	Repeat/repeat + rephrase + new question	-60	15	.02	16	12	0	17	13	.50
138.	Rephrase/repeat + rephrase + new question	-03	15	.92	16	13	.19	17	14	.49
139.	Brief feedback/brief + long feedback	49	18	.04			NONE			NONE
R.	Math Contacts									
140.	Total public math contacts/total public math contacts + total private math contacts	00	00	1.00		NO DATA			NO DATA	
141.	Total teacher initiated private math contacts/ total public math contacts + total private math contacts	00	00	1.00		NO DATA			NO DATA	

Table 7 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I N	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II N	YEAR II MODE

[1] Process variables for which data were collected in the first and second years of the study came from an expanded version of the Brophy-Good Dyadic Observational System. Variables are either proportions or frequencies from the different interaction categories. These data were taken during afternoon, whole-class observations.

NV No meaningful variance (only 0, 1, or 2 non-modal scores)

* Significantly skewed: skewness $p < .05$

NONE No one mode: Bi- or multimodal

Table 8
STABILITY CORRELATIONS AND DISTRIBUTION DATA
FOR TEACHER PROCESS VARIABLES FROM THE
EXPANDED BROPHY-GOOD DYADIC OBSERVATIONAL SYSTEM: READING GROUP OBSERVATIONS[1]

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
A.	Selecting Respondents to Questions									
1.	% Preselects respondent before asking questions	-30	15	.27	17	10	0*	17	12	0*
2.	Calls on non-volunteer	69	15	.00			NONE			NONE
3.	Calls on volunteer	70	15	.00			NONE	17	14	.15
4.	Student calls out answers	53	15	.04			NONE	17	13	.03
B.	Difficulty Level of Questions									
5.	Process questions/process + product questions	19	15	.51	17	13	0	17	13	0
6.	Choice questions/process + product + choice	-16	15	.57	17	12	0*			NONE
C.	Quality of Children's Answers									
7.	% Correct	25	15	.37			NONE*			NONE
8.	% part-correct	20	15	.49	17	14	0*	17	14	.02*
9.	% wrong	07	14	.80	17	14	.06*	17	14	.10
10.	% "don't know"	-26	15	.35	17	6	0*			NONE*
11.	% no response	33	15	.22			NONE*			NONE

Table 8 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II MODE
D.	Teacher Reactions to Correct Answers									
12.	Praise	-02	15	.96	17	8	0			NONE*
13.	Criticism	03	15	.92			NV	17	3	0*
14.	Failure to give feedback	-25	15	.38	17	5	0*	17	8	0*
15.	Process feedback	03	15	.92	17	3	0*	17	10	0*
16.	New question	00	11	1.00			NONE			NONE*
E.	Teacher Reactions to Part-Correct Answers									
17.	Praise	00	15	1.00			NV			NV
18.	Criticism	00	11	1.00			NV			NV
19.	Failure to give feedback	00	11	1.00			NV			NV
20.	Process feedback	00	11	1.00			NV			NV
21.	Gives the answer	48	11	.06	14	6	0*	16	9	0
22.	Calls on someone else	48	11	.14	14	8	0	16	11	0
23.	Another student calls out the answer	00	11	1.00			NV	16	5	0*
24.	Repeats, rephrases, or asks new question	-24	11	.48	14	10	0			NONE
25.	Repeats question	59	11	.05	14	5	0*	16	12	0

Table 8 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE	YEAR I MODE	YEAR II N	YEAR II N-MODE	YEAR II MODE
26.	Rephrases or gives clue	30	11	.37	14	8	0	16	13	.01
27.	Asks new question	01	11	.98	14	7	0	16	7	0*
F.	Teacher Reactions to Wrong Answers									
28.	Praise	00	15	1.00			NV			NV
29.	Criticism	-18	15	.52			NV	17	5	0*
30.	Failure to give feedback	-12	15	.68			NV	17	4	0*
31.	Process feedback	04	15	.88			NV	17	9	0*
32.	Gives the answer	12	15	.68	16	12	0	17	14	.25
33.	Calls on someone else	15	15	.61			NONE	17	14	.38
34.	Another student calls out the answer	-01	15	.97	16	5	0*	17	8	0*
35.	Repeats, rephrases, or asks new question	21	15	.46			NONE			NONE*
36.	Repeats question	-04	15	.87	16	8	0			NONE
37.	Rephrases or gives clue	06	15	.82	16	11	0*	17	13	.25
38.	Asks new question	-06	15	.84	16	7	0	17	6	0*

Table 8 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
G.	Teacher Reactions to "I Don't Know" or No Response									
39.	Criticism	00	13	1.00			NV	17	3	0*
40.	Failure to give feedback	-23	13	.46			NV	17	5	0*
41.	Gives the answer	-15	13	.62	15	5	0*	17	13	0
42.	Calls on someone else	-23	13	.45			NONE	17	14	.63
43.	Another student calls out the answer	-09	13	.77	15	4	0*	17	8	0*
44.	Repeats, rephrases, or asks new question	00	13	.99			NONE	17	13	0
45.	Repeats question	-57	13	.04	15	8	0*	17	12	0
46.	Rephrases or gives clue	-01	13	.97	15	10	0*	17	12	0
47.	Asks new question	04	13	.90	15	5	0*	17	4	0*
H.	Teacher Reactions Combined Across All Response Opportunities									
48.	Praise	02	15	.95	17	8	0			NONE*
49.	Failure to give feedback	-20	15	.49	17	7	0*	17	10	0*
50.	Process feedback	20	15	.49	17	3	0*	17	13	0*

Table 8 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
51.	New question	08	15	.76	17	14	.09			NONE*
52.	Repeat, rephrase, or new question after failure to answer	50	15	.06	17	14	.33	17	15	.39
53.	Repeats question after failure to answer correctly	35	15	.20	17	11	0*	17	14	.03*
54.	Gives the answer after failure to answer correctly	07	15	.80	17	14	0*			NONE
55.	Calls on another student after failure to answer correctly	66	15	.01	17	14	.09			NONE
56.	Another student calls out answer after failure to answer correctly	-37	15	.18	17	7	0*	17	12	0*
I.	Student Response Opportunities									
57.	Response opportunities/ total teaching time	.75	15	.00			NONE*			NONE
J.	Student Initiated Questions (SIQ's)									
58.	% SIQ's irrelevant	03	10	.93	13	6	.01*	14	3	0*
59.	% SIQ's called out	-13	10	.71	13	8	.02	14	9	.98
60.	Praise of question after relevant SIQ	00	9	1.00			NV			NV

Table 8 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
61.	Criticism of question after relevant SIQ	00	9	1.00			NV			NV
62.	% relevant SIQ's delayed	-33	9	.39	12	3	0*	14	4	0
63.	% relevant SIQ's not accepted	96	9	.00	12	4	0*			NV
64.	% relevant SIQ's given brief feedback	-16	9	.68	12	8	.49			NONE*
65.	% relevant SIQ's given long feedback	-23	9	.56	12	7	0*	14	8	0*
66.	% relevant SIQ's redirected to class	00	9	1.00			NV	14	3	0*
67.	Behavioral praise of relevant SIQ	00	9	1.00			NV			NV
68.	Behavioral criticism of relevant SIQ	00	9	1.00			NV			NV
69.	Behavioral warning after relevant SIQ	00	9	1.00			NV			NV
70.	Criticism of question after irrelevant SIQ	00	2	1.00			NV			NV
71.	% irrelevant SIQ given no feedback	00	2	1.00			NV			NV
72.	% irrelevant SIQ delayed	00	2	1.00			NV			NV
73.	% irrelevant SIQ given brief feedback	1.0	2	1.00			NV			NONE
74.	% irrelevant SIQ given long feedback	1.0	2	1.00			NV			NONE
75.	% irrelevant SIQ not accepted	00	2	1.00			NV			NV

Table 8 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
76.	% irrelevant SIQ redirected to class	00	2	1.00			NV			NV
77.	Behavioral criticism after irrelevant SIQ	00	2	1.00			NV			NV
78.	Behavioral warning after irrelevant SIQ	00	2	1.00			NV			NV
K.	Student Initiated Public Interactions									
79.	Student initiated comments and questions/total response opportunities	21	15	.45	17	13	0*	17	13	.02*
L.	Student Initiated Comments (SIC's)									
80.	% SIC's relevant	-36	12	.25	14	8	1.00*	16	10	1.00
81.	% SIC's called out	05	12	.88	14	9	0	16	12	1.00
82.	Praise of comment after relevant SIC	-26	11	.45	13	5	0*	16	7	0*
83.	% relevant SIC's given no feedback	06	11	.84	13	5	0	16	5	0*
84.	% relevant SIC's delayed	04	11	.89	13	3	0*			NV
85.	% relevant SIC's not accepted	-34	11	.31	13	4	0*	16	10	0*
86.	% relevant SIC's accepted	04	11	.91	13	10	.50	16	12	.75
87.	% relevant SIC's integrated into discussion topic	25	11	.46	13	7	0*	16	9	0

Table 8 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
88.	% relevant SIC's which cause a shift in topic	00	11	1.00	13	3	0*			NV
89.	Behavioral praise after relevant SIC's	00	11	1.00			NV			NV
90.	Behavioral criticism after relevant SIC's	-14	11	.68	13	4	0*	16	3	0*
91.	Behavioral warning after relevant SIC's	03	11	.93			NV	16	5	0*
92.	Praise of comment after irrelevant SIC's	00	4	1.00			NV			NV
93.	% irrelevant SIC's given no feedback	00	4	1.00			NV	10	3	0*
94.	% irrelevant SIC's delayed	1.00	4	.00			NV			NV
95.	% irrelevant SIC's not accepted	-56	4	.45			NV	10	3	0
96.	% irrelevant SIC's accepted	02	4	.98	7	3	1.00			NONE
97.	% irrelevant SIC's integrated into discussion topic	00	4	1.00			NV			NV
98.	% irrelevant SIC's which cause a shift in topic	00	4	1.00			NV			NV
99.	Behavioral criticism after irrelevant SIC's	00	4	1.00			NV			NV
100.	Behavioral warning after irrelevant SIC's	00	4	1.00			NV			NV
M.	Self and Opinion Questions									
101.	Self questions/process + product + choice questions	-16	14	.58	16	9	0	17	13	.02*

Table 8 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
102.	% self questions which were subject-matter related	00	1	1.00			NV			NONE
103.	% self questions related to personal preference	00	1	1.00			NV	14	12	.02
104.	Opinion questions/process + product + choice questions	39	13	.19	15	9	0*	17	13	0*
105.	% opinion questions given no response	00	07	1.00	9	3	0*	13	3	0*
106.	% opinion questions followed by praise	00	07	1.00			NV	13	5	0*
107.	% of opinion questions followed by teacher disagreement	-12	7	.79			NV	13	3	0*
108.	% student opinions accepted	-65	7	11	9	6	1.00	13	9	1.00
109.	% student opinions integrated into discussion topic	-27	7	.56	9	4	0			NV
N.	Private Dyadic Contacts									
110.	% private contacts student initiated	38	11	.25			NONE			NONE
111.	Student initiated work contacts involving praise	-07	10	.85	12	5	0*	17	11	0
112.	Student initiated work contacts involving criticism	00	10	1.00			NV	17	9	0*

Table 8 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
113.	% of private work contacts student initiated	-20	10	.59	12	10	.63			NONE
114.	% student initiated contacts delayed	55	10	.10	12	6	0	17	12	0*
115.	% student initiated contacts given brief feedback	11	10	.75	12	10	.28			NONE
116.	% student initiated contacts given long feedback	18	10	.61			NONE			NONE*
117.	% student initiated contacts involving personal concerns	18	10	.62	12	9	.20			NONE
118.	% student initiated requests delayed	-12	9	.76	11	5	0	17	10	0
119.	Private work contacts/private work contacts + public response opportunities	69	12	.01			NONE	17	13	.30
120.	Procedural contacts/procedural contacts + response opportunities	67	12	.02			NONE	17	14	.08
121.	Teacher initiated work contacts/teacher initiated work + procedure contacts	41	11	.21			NONE			NONE
122.	Teacher initiated work contacts involving praise	-06	11	.86	12	5	0	17	13	0*
123.	Teacher initiated work contacts involving mere observation	93	11	.00			NV	17	8	0*

Table 8 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
124.	Teacher initiated work contacts involving brief feedback	08	11	.82	12	8	1.00			NONE
125.	Teacher initiated work contacts involving long feedback	-17	11	.62	12	8	0	17	14	.15
126.	% teacher initiated procedural contacts which were management requests	-01	11	.98			NONE			NONE*
127.	% teacher thanks student for doing a favor request	-24	8	.57			NV	14	6	0*
128.	% teacher thanks student for following a management request	-10	11	.77	13	3	0*	17	10	0
O.	Combined Teacher Evaluation Statements									
129.	Academic praise/academic praise + academic criticism	53	10	.11	12	7	1.00*	17	13	.99*
130.	Behavioral praise/total behavioral contacts	-01	10	.97			NV	17	7	0*
131.	Behavioral warnings/behavioral warnings + behavioral criticism	06	9	.86	11	8	.51			NONE
P.	Discipline and Control Errors									
132.	% Discipline contacts involving one or more errors	58	10	.07	12	7	0*	17	8	0*

Table 8 (Continued)

NO.	VARIABLE NAME	r	N	p*	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE
133.	Target errors/total errors	00	5	1.00			NV	8	4	0*
134.	Timing errors/total errors	-36	5	.55	7	3	0	8	3	.01
135.	Overreactions/total errors	86	5	.06			NV	8	5	.99
136.	Nonverbal control contacts/total control contacts	22	10	.55	12	7	0*	17	13	0
Q.	Combined Teacher Feedback Data									
137.	Repeat/repeat + rephrase + new question	22	14	.45	16	12	0			NONE
138.	Rephrase/repeat + rephrase + new question	49	14	.07	16	13	.34	17	14	.28
139.	Brief feedback/brief + long feedback	-14	11	.67	13	11	.88	17	15	.88
R.	Math Contacts									
140.	Total public math contacts/total public math contacts + total private math contacts	00	00	1.00		NO DATA			NO DATA	
141.	Total teacher initiated private math contacts/ total public math contacts + total private math contacts	00	00	1.00		NO DATA			NO DATA	

Table 8 (Continued)

NO.	VARIABLE NAME	r	N	p	YEAR I N	YEAR I N-MODE N	YEAR I MODE	YEAR II N	YEAR II N-MODE N	YEAR II MODE

[1] Process variables for which data were collected in the first and second years of the study came from an expanded version of the Brophy-Good Dyadic Observational System. Variables are either proportions or frequencies from the different interaction categories. These data were taken during small group (reading group) observations.

NV No meaningful variance (only 0, 1, or 2 non-modal scores)

[a] Significantly skewed: skewness $p < .05$

NONE No one mode: Bi- or Multimodal